INTER-AMERICAN TROPICAL TUNA COMMISSION
COMISIÓN INTERAMERICANA DEL ATÚN TROPICAL

SCIENTIFIC WORKING GROUP

4TH MEETING

LA JOLLA, CALIFORNIA (USA)
19-21 MAY 2003

REPORT ON SAMPLING THE EASTERN PACIFIC OCEAN TUNA CATCH FOR SPECIES COMPOSITION AND LENGTH-FREQUENCY DISTRIBUTIONS

by

Patrick K. Tomlinson

1. INTRODUCTION

The area of concern to the Inter-American Tropical Tuna Commission, referred to in this report as the eastern Pacific Ocean (EPO), is defined as that between 40 degrees North and 40 degrees South and from 150 degrees West to the western edge of the North and South American continents. In Background Paper A21, prepared for the meeting of the Scientific Working Group convened in La Jolla, California, USA, on April 10-14, 2000, a model to be used for simultaneous estimation of the surface fishery catch of individual species of tunas from the EPO, and also their length-frequency distributions, was presented. Appropriate parts of that paper are included as Appendix I in this document for completeness. In this document, we will report on the implementation of that model in the EPO during the years 2000, 2001, and 2002.

The catch of the EPO tuna surface fishery consists of five major tuna species, plus some other minor tunas and tuna-like species. Of the five major species, only the catches of yellowfin, skipjack, and bigeye tuna are considered, with some reference to bluefin tuna, but nothing about albacore tuna. The EPO surface fishery is divided into 12 calendar months, 13 areas (Figure 1), and the following 7 fishing methods:

1. baitboats (includes also bolicheras and jigboats);
2. small purse seiners (<364 metric tons (mt)) setting on schools associated with floating objects;
3. small purse seiners setting on schools not associated with anything;
4. small purse seiners setting on schools associated with dolphins;
5. large purse seiners (=364 mt) setting on schools associated with floating objects;
6. large purse seiners setting on schools not associated with anything;
7. large purse seiners setting on schools associated with dolphins.

This results in 1092 categories, which will be referred to as strata. The estimation procedure of a hypothetical stratum is given in Appendix 1. To estimate for the monthly or annual totals, the procedure is to estimate for each stratum separately, and then sum the results for all of the appropriate strata together.

There are three different types of sampling. First, the sampler selects a vessel that is being unloaded and determines if any wells being unloaded have fish from only one stratum. Wells containing a mixture of tunas from different strata are not sampled. Second, the sampler is to count, independently from measuring, the number of each species in a random sample of several hundred fish. Third, the sampler randomly removes a number of fish of each species (usually 50) and measures and records the lengths of each to the nearest millimeter. In addition to the sampled data, an estimate of the total catch (yellowfin + skipjack + bigeye) for each stratum is needed. There have been exceptions to these rules, which will be discussed later. For a more complete explanation of the model, see Appendix 1.

2. EXCEPTIONS TO THE SAMPLING MODEL

There were 903, 1091, and 916 wells sampled during the years 2000, 2001, and 2002 respectively (Table 1). There may be a few samples that have not been processed yet. Also, information provided by vessels (logbooks), dolphin observer records, and cannery reports are all important in determining the catch by stratum, but some of these are incomplete. Therefore, this report should be considered as preliminary. Of the sampled wells, species counts were made for 780, 875, and 816, respectively. Other wells were assigned estimates of species composition and size frequency based on other criteria, described below. Of the 439 wells without counts, 208 came from wells containing a single species and are therefore equivalent to samples with species counts. This leaves 231 wells that required a different treatment. Also, 676 of the 2909 wells came from vessels that sorted the fish by size and/or species before unloading (hereafter referred to as a split well), making it impossible to obtain either a random species count from the well or a random sample of each species for measuring. The treatment of these is again described below, except to point out that the vessel's estimates of the species composition in weight, by species, for the sampled well was utilized.

Split wells were treated in such a way as to produce a sample from each such well that has the same data organization as the samples that were obtained from the non-split wells. This involves first estimating the number of fish of each species in each of the splits by sampling a number of fish for size, computing the average weight, and dividing into the split's total weight provided by the vessel. Summing these estimates provides an estimate of the species composition of the well that is used in place of the random count. Using this species-composition estimate, the fish that were measured are resampled with probability proportional to the composition estimate. These resampled measurement data are then substituted for the original measurement data. This provides a data set that has the same structure as the non-split wells. This facilitates the use of the estimation algorithms given in Appendix 1.

Finally, there is the problem of strata with catches, but no samples, which exists primarily for strata with catches of less than 1000 mt per month. Based on previous experience and the belief that data for other areas or gears or months can be used to represent these strata without samples, data from sampled strata are copied for use as samples from the unsampled strata. This provides data bases that can be used to obtain the estimates for every stratum with catch.

3. THE STANDARD METHOD VERSUS THE SPECIES-COMPOSITION METHOD

The standard method, which has been used for many years, involves treating each species as a separate sampling problem (see Tomlinson *et al.*, 1992), and in many cases only one species was measured when there may have been more than one species in the well. Therefore, it is not possible to apply the species-composition method to data for past years. It is possible, however, to use the data collected from 2000 to 2002 as if it were collected by the standard method. The main differences between the two methods lie in the construction of a table showing the catch by stratum. For the standard method, the vessel's logbook data, dolphin observer data, and cannery statistics are treated as being correct with respect to species, and a separate estimate is made for each species for each stratum. For the species-composition method, the species are added together in each of these data sets and treated as total catch of yellowfin, skipjack, and bigeye combined before estimating the total catch by stratum, and then the sampling data are used to separate these stratum totals into catch by species.

Applying both methods to their respective data sets, estimating catch by stratum by species, and then summing strata within months and then across months provides annual estimates by species that can be compared (Table 2, Figure 2).

The annual length-frequency distributions, by species, obtained by the standard method and by the species composition method are similar(Figures 3a, 3b, 3c). The greatest differences are for bigeye, with the species-composition method estimating more fish for most length groups, especially in 2000 in the 90- to 110-cm interval, in 2001 in the 40- to 100-cm interval, in 2001 in the 115- to 150-cm interval, and in 2002 in the 40- to 150-cm interval. For a few lengths, the species composition method estimated fewer than the standard method.

4. ESTIMATION OF ERROR IN THE ESTIMATES

For the standard method, the statistics on total catch by species do not involve sampling. The standard method depends on the buyers or processors providing statistics on the weight of each species purchased. However, availability of these data lags behind the time of unloading the catches, and some of the values used were based on data provided by other sources. Thus, the values given in this paper are subject to change.

With the species-composition method, the same type of lag in the final data exists, but for the combined catch it is not as serious since it is not nearly as difficult to estimate the total catch on board a vessel as it is to estimate the amounts by species, as required by the standard method. Since there are three sources of variation that can be considered for the species-composition method (differences among wells, differences in sizes of fish, and differences in the species composition), it is possible to construct an estimate of the error associated with the estimates. We will discuss only the estimation of error associated with the annual estimates of catch in metric tons for yellowfin, skipjack, and bigeye, since the species composition method has not been applied to bluefin.

The assumptions utilized for the estimation are:

1. The wells sampled within a stratum (or the substitutes mentioned above) are a simple random sample of the wells from that stratum and can be resampled at random, with replacement, to obtain among-well variance.
2. The fish sampled for measurement are a simple random sample from the well, and provide estimates of the average size, and its standard deviation, for each species present. These means and standard deviations can be used to resample from a population with those means and standard deviations. This, together with step 3, provides estimates of the variation in number, by species, within the well.
3. The fish counted within a well are a simple random sample with a trinomial distribution. Assuming that the trinomial estimated is the population trinomial, resampling from the distribution provides an estimate of the within-well variance in species composition.

Resampling the sample data N times provides N estimates of the catch by species. These N estimates can be used to provide a mean and standard deviation of the N estimates. Experience from past estimation (see Tomlinson, 2002) showed that these distributions are reasonably normally distributed. Therefore, the standard deviations obtained from the resampling were used to obtain 95 percent confidence regions of the estimates (Table 2, Figure 2).

This resampling technique (see Tomlinson, 2002) was also used to determine the importance of each of the three types of sampling (wells, average weight, counts). The results showed that most of the variance comes from among wells and that the three sources of error are probably independent of each other. Reductions in standard deviations could best be obtained by increasing the number of wells, while retaining the same number of fish counted and measured. In theory, the number counted and measured could be reduced, but there is an assumption of randomness where no rigorous random procedure exists, and the larger numbers help to ensure that this assumption is met. Because of the high proportion of wells that contain fish from more than one stratum and the cost of placing samplers at more unloading sites, it is impractical to sample more wells. Besides, the standard deviations obtained give coefficients of variation of less than 10 percent (which is often used as a criterion). However, there are still many strata with catches of less than 1000 tons that are not sampled.

5. LITERATURE CITED

Tomlinson, Patrick K., Sachiko Tsuji, and Thomas P. Calkins. 1992. Length-frequency estimation for yellowfin tuna (*Thunnus albacares*) caught by commercial fishing gear in the eastern Pacific Ocean. Inter-Amer. Trop. Tuna Comm., Bull., 20 (6): 357-398.

Tomlinson, Patrick K. 2002. Progress on sampling the eastern Pacific Ocean tuna catch for species composition and length-frequency distributions. Inter-Amer. Trop. Tuna Comm., Stock Assessment Report 2: 339-365.

APPENDIX

SAMPLING THE CATCH SIMULTANEOUSLY FOR SPECIES COMPOSITION AND LENGTH FREQUENCIES IN THE MULTI-SPECIES SURFACE FISHERY FOR TUNAS OF THE EASTERN PACIFIC OCEAN

DEFINITION OF THE POPULATION AND THE SAMPLING MODEL

Given the 13 areas (Figure 1), 12 months, and 7 fishing methods, there is a total of 1,092 strata for each year. The population is defined as the annual catch for all strata combined. Summaries across areas, months, or gear categories, or any combination of these, are obtained by summing the estimates for the appropriate strata. Therefore, it is only necessary to define the sampling model for a single stratum and redefine the population as the catch within one stratum.

Within each stratum with catch, the sampling is done during the unloading of a vessel. Each vessel has several compartments, called wells. These wells of fish are the first stage, and the individual fish within the well are the second stage, in a stratified two-stage sampling scheme. Unfortunately, neither the number of wells in a stratum nor the number of fish in a well will be known in advance. Therefore, the wells to be sampled are selected as time and availability permit, and it will be assumed that this is equivalent to a simple random sample of the wells.

1. There are several possible sources of error in the assumption of a simple random sample of wells. The first, which has been mentioned, prohibits truly randomizing the selection process. Associated with the assumption of randomness of wells selected is the assumption that wells within the same vessel being unloaded are independent of each other. In some cases, at least, this will not be true. At times, the fish from more than one well will be unloaded simultaneously, so the samples will come from more than one well. In these cases the multiple wells will be treated as a single well.
2. Many wells filled with purse seine-caught fish contain fish that were caught in more than one month, or in more than one area, or by more than one gear category (fish caught in sets of different types loaded into the same well). These wells must be excluded from the sampling. It is necessary to assume that this does not introduce a problem. In other words, the assumption is that wells filled with a single set type (of the possible three) are representative of all sets of the same type made within the stratum.

Some of the problems associated with using wells as the first-stage unit could be solved if sets were used as the first stage, but, since the sampling cannot be done at sea, it is not possible to do this. Within each well sampled, there will be one or more species, and a good approximation of the total weight of all fish within the well will be available. Approximations of the weight of fish of each species, which might be available, will not be considered as usable values. Until experience dictates otherwise, it will be assumed that the entire catch will consist of one to three species of tuna (yellowfin, skipjack, and/or bigeye). Most strata will have a least two species and the others will have all three or only one. The fish to be sampled from the well are to be selected one at a time as circumstances permit. As with the first-stage units, it is not possible to guarantee a truly random selection of fish, so randomness will be assumed by assuming that the order of unloading the fish is random.

Define

S = number of species;

Q = number of wells landed in the stratum;

N_{ij} = number of fish of species i in well j;.. (1)

$N_j = \sum_i^S N_{ij}$ = total number of fish of all species in well j;.. (2)

$N_i = \sum_{j}^{Q} N_{ij}$ = total number of fish of species i in the stratum; .... (3)

$N = \sum_{i}^{S} \sum_{j}^{Q} N_{ij}$ = total number of fish of all species landed in the stratum; .... (4)

W_{ij} = weight of fish of species i in well j; .... (5)

$W_j = \sum_{i}^{S} W_{ij}$ = total weight of fish of all species in well j; .... (6)

$W_i = \sum_{j}^{Q} W_{ij}$ = total weight of species i in the stratum; .... (7)

$W = \sum_{i}^{S} \sum_{j}^{Q} W_{ij}$ = total weight of fish of all species landed in the stratum; .... (8)

$\overline{W} = W \div N$ = average weight of all species combined in the stratum; .... (9)

$\overline{W}_i = W_i \div N_i$ = average weight of species i in the stratum; .... (10)

$\overline{W}_j = W_j \div N_j$ = average weight of all species combined in well j; .... (11)

$\overline{W}_{ij} = W_{ij} \div N_{ij}$ = average weight of species i in well j; and .... (12)

$F_{ij} = N_{ij} \div N_j$ = fraction of catch in number of fish of species i in well j. .... (13)

Since

$\overline{W}_{ij} F_{ij} = (W_{ij} \div N_{ij})\,(N_{ij} \div N_j) = W_{ij} \div N_j$ .... (14)

then

$\sum_{i}^{S} \overline{W}_{ij} F_{ij} = \sum_{i}^{S} (W_{ij} \div N_j) = (1 \div N_j) \sum_{i}^{S} W_{ij} = \overline{W}_j$ .... (15)

If

$P_{ij} = W_{ij} \div W_j$ = fraction of catch of species i in well j,

then

$P_{ij} = (\overline{W}_{ij} N_{ij}) \div (\overline{W}_j N_j) = (\overline{W}_{ij} F_{ij}) \div \overline{W}_j$ .... (16)

Continuing,

$F_i = N_i \div N$ = fraction of total stratum catch, in numbers, of species i; .... (17)

$P_i = W_i \div W$ = fraction of total stratum catch, in weight, of species i; .... (18)

L_{ijh} = length of the hth fish of species i in well j;

w_{ijh} = weight of the hth fish of species i in well j; and

$w_{ijh} = aL_{ijh}^{\,b}$ = estimate of w_{ijh}, where a and b are parameters of the weight-length relationship.

Assume that $W_{ij} = \sum_{h}^{Nij} aL_{ijh}^{\ b}$ ....(19)

T = number of length groups, each group encompassing 1 cm of length;

N_{ijk} = number of fish in the kth length group of species i in well j; and ....(20)

$F_{ijk} = N_{ijk} \div N_{ij}$ = fraction of species i, in number, in well j that belong to length group k.................(21)

Then,

$$N_{ij} = \sum_{k}^{T} N_{ijk} \quad \text{......(22)}$$

Note: The value of k is found by measuring the fish to the nearest millimeter and then truncating the length measurement to centimeters. For example, a 431-mm fish would belong to length group 43 and an 1132-mm fish to group 113.

Then,

W_{ijk} = total weight of the N_{ijk} fish; ....(23)

$W_{ij} = \sum_{k}^{T} W_{ijk}$; and..(24)

$P_{ijk} = W_{ijk} \div W_{ij}$ = fraction of species i, in weight, in well j that belong to length group k.................(25)

Then,

$N_{ik} = \sum_{j}^{Q} N_{ijk}$ = total number of species i in length group k in the stratum;....................................(26)

$W_{ik} = \sum_{j}^{Q} W_{ijk}$ = total weight of species i in length group k in the stratum;....................................(27)

$F_{ik} = N_{ik} \div N_i$ = fraction of catch, in numbers, of species i in length group k in the stratum; and ....(28)

$P_{ik} = W_{ik} \div W_i$ = fraction of the catch, in weight, of species i in length group k in the stratum...........(29)

OBJECTIVES

There are two primary objectives to be accomplished through sampling of the catch. The first is to obtain an estimate of Equation 7 (species composition by weight), and the second is to obtain an estimate of Equation 26 (catch in numbers by length group). The secondary objectives are to obtain an estimate of Equation 10 (average weight by species) and of Equation 27 (catch in weight by length group).

For each well which enters the sampling, it will be assumed that the W_j (Equation 6) is known and for each stratum, W (Equation 8) is known. Also, for each sampled well, the sample data will allow estimation of $\overline{W}_{ij}$ (Equation 12) and F_{ij} (Equation 13). These two estimates are used to estimate $\overline{W}_{j}$ (Equation 15) and $N_j = W_j \div \overline{W}_j$ (Equation 2). The sampling data can also be used to estimate F_{ijk} (Equation 21) and P_{ij} (Equation 16).

WITHIN-WELL ESTIMATION

First, a well from a single stratum is chosen at random, without replacement. Two independent sampling schemes are carried out as the fish are unloaded from the chosen well. The first consists of counting a fixed number (Equation 36) of fish (approximately 400 will be counted to begin the program) at random, without replacement, identifying the species and recording the numbers of fish of each species (Equation 30). The second consists of removing approximately 50 fish (Equation 37) at random without replacement foreach species observed and recording the species and length in millimeters (Equation 31) of each of these. An estimate of the total catch in weight of all species combined (Equation 6) in the sampled well is also recorded and treated as if it were the exact weight.

Let

n_{ij} = number of fish of species i from well j recorded during the counting sample; ....(30)

l_{ijh} = length of the hth fish of species i measured from well j; and ....(31)

$w_{ijh} = al_{ijh}^{b}$ = estimated weight of the hth fish of species i measured from well j. ....(32)

Let k = integer part of ($l_{ijh} \div 10$) ....(33)

m_{ijk} = number of fish measured of length group k of species i in well j; ....(34)

w_{ijk} = total weight of the m_{ijk} fish based on the w_{ijh}; and ....(35)

$n_j = \sum_{i}^{S} n_{ij}$ = number of fish counted from well j ....(36)

Let $k_{\max}$ be the length group of the longest fish.

$m_{ij} = \sum_{k=1}^{k\max} m_{ijk}$ = number of fish of species i measured from well j; ....(37)

$w_{ij} = \sum_{k=1}^{k\max} w_{ijk}$ = weight of fish of species i measured from well j; ....(38)

$\hat{f}_{ij} = n_{ij} \div n_j$ = estimate of F_{ij} (Equation 13); ....(39)

$\hat{f}_{ijk} = m_{ijk} \div m_{ij}$ = estimate of F_{ijk} (Equation 21); ....(40)

$\bar{w}_{ij} = w_{ij} \div m_{ij}$ = estimate of $\bar{W}_{ij}$ (Equation 12); ....(41)

$\bar{w}_{j} = \sum_{i}^{S} \bar{w}_{ij} \hat{f}_{ij}$ = estimate of $\bar{W}_{j}$ (Equations 11 and 15); ....(42)

$\bar{w}_{i} = \sum_{j}^{q} \hat{W}_{ij} \div \sum_{j}^{q} \hat{N}_{ij}$ = estimate of $\bar{W}_{i}$ (Equation 10); ....(43)

$\hat{N}_{j} = W_j \div \bar{w}_{j}$ = estimate of N_j (Equation 2); ....(44)

$\hat{N}_{ij} = \hat{N}_{j} \hat{f}_{ij}$ = estimate of N_{ij} (Equation 1); ....(45)

$\hat{N}_{ijk} = \hat{N}_{ij} \hat{f}_{ijk}$ = estimate of N_{ijk} (Equation 20); ....(46)

$\hat{p}_{ij} = \bar{w}_{ij} \hat{f}_{ij} \div \bar{w}_{j}$ = estimate of P_{ij} (Equation 16); ....(47)

$\hat{p}_{ijk} = w_{ijk} \div w_{ij}$ = estimate of P_{ijk} (Equation 25); ....(48)

$\hat{W}_{ij} = W_j \hat{p}_{ij}$ = estimate of W_{ij} (Equation 5); and ....(49)

$\hat{W}_{ijk} = \hat{W}_{ij} \hat{p}_{ijk}$ = estimate of W_{ijk} (Equation 23). ....(50)

WITHIN-STRATUM ESTIMATION

Let

q = number of wells sampled from a single stratum;

$\hat{f}_{ik} = \sum_j^q \hat{N}_{ijk} \div \sum_j^q \hat{N}_{ij}$ = estimate of F_{ik} (Equation 28); ....(51)

$\hat{p}_{ik} = \sum_j^q \hat{W}_{ijk} \div \sum_j^q \hat{W}_{ij}$ = estimate of P_{ik} (Equation 29); ....(52)

$\overline{w} = \sum_j^q W_j \div \sum_j^q \hat{N}_j$ = estimate of $\overline{W}$ (Equation 9); ....(53)

$\hat{f}_i = \sum_j^q \hat{N}_{ij} \div \sum_j^q \hat{N}_j$ = estimate of F_i (Equation 17); ....(54)

$\hat{p}_i = \sum_j^q \hat{W}_{ij} \div \sum_j^q \hat{W}_j$ = estimate of P_i (Equation 18); ....(55)

$\hat{N} = W \div \overline{w}$ = estimate of N (Equation 4); ....(56)

$\hat{N}_i = \hat{N} \hat{f}_i$ = estimate of N_i (Equation 3); ....(57)

$\hat{W}_i = W \hat{p}_i$ = estimate of W_i (Equation 7); ....(58)

$\hat{N}_{ik} = \hat{N}_i \hat{f}_{ik}$ = estimate of N_{ik} (Equation 26); and ....(59)

$\hat{W}_{ik} = \hat{W}_i \hat{p}_{ik}$ = estimate of W_{ik} (Equation 27). ....(60)

Equations 58 and 59 satisfy the two primary objectives, and Equations 43 and 60 satisfy the two secondary objectives.

FIGURE 1. Areas used for sampling lengths of surface-caught tunas in the eastern Pacific Ocean (EPO)

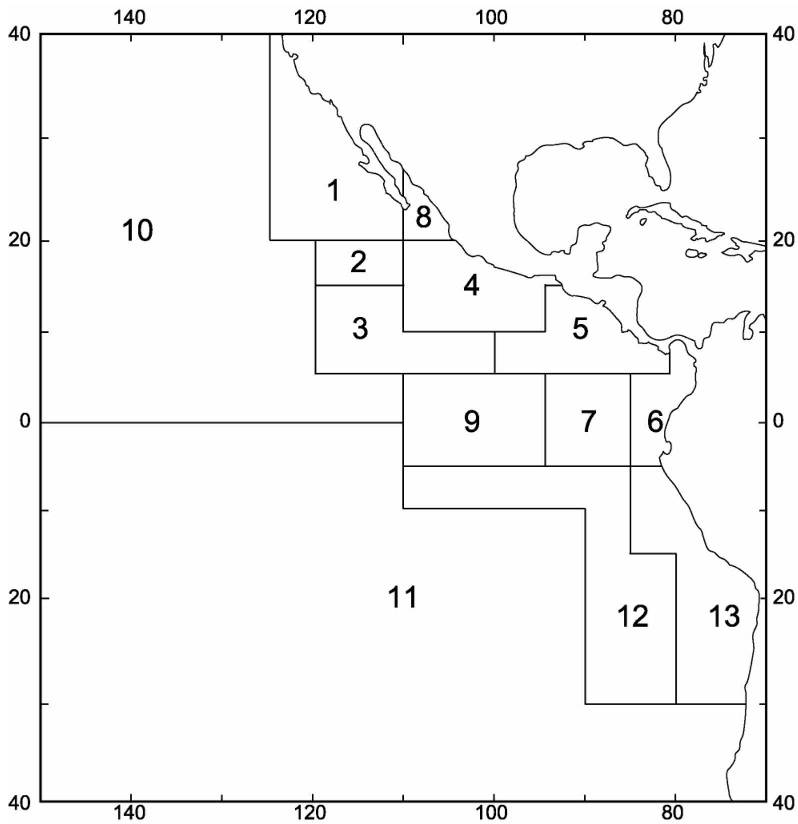

FIGURE 2. Catches from the standard versus the composition method.

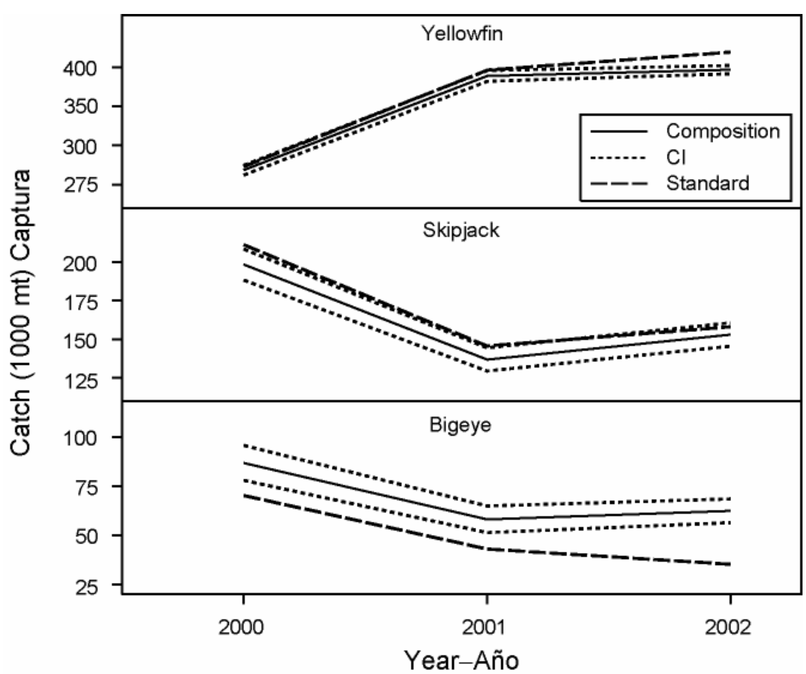

FIGURE 3a. Yellowfin surface fishery catches, standard versus composition method.

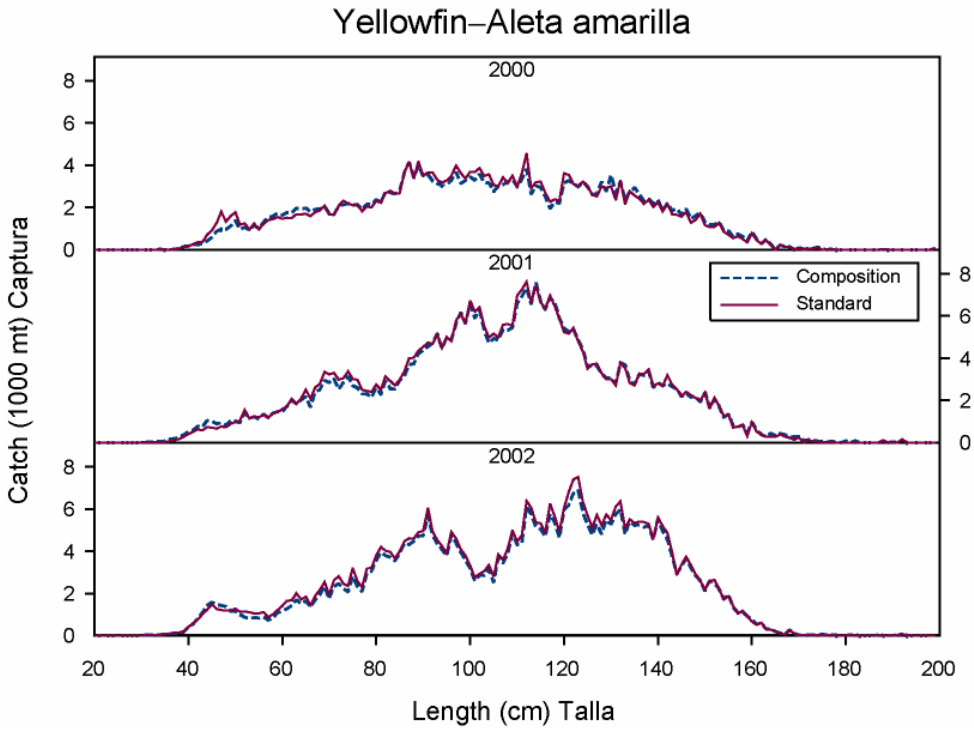

FIGURE 3b. Skipjack surface fishery catches, standard versus composition method..

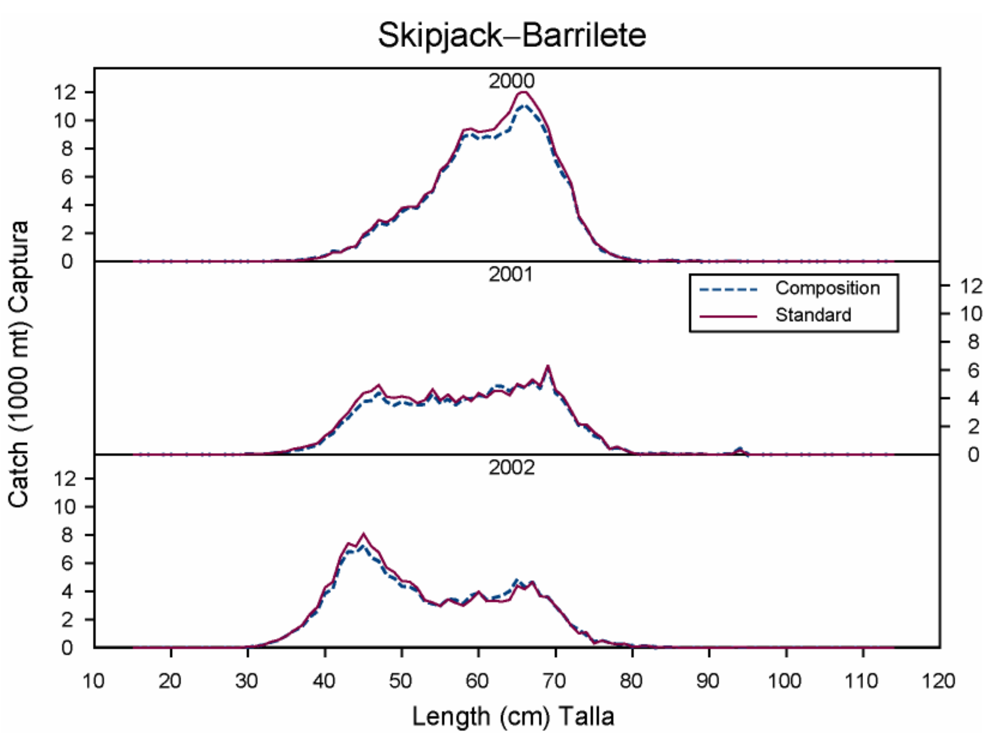

FIGURE 3c. Bigeye surface fishery catches, standard versus composition method..

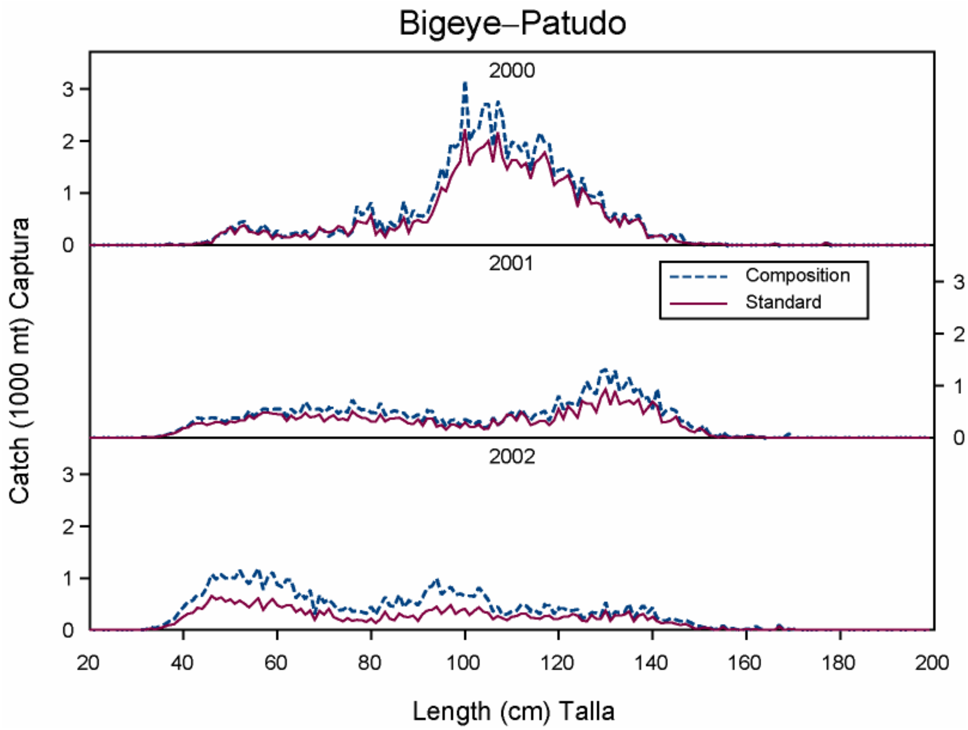

TABLE 1. Number of wells sampled and number of fish measured, 2000-2002.

	2000	**2001**	**2002**
Number of wells sampled:			
Non-split wells	769	829	635
Split wells	134	262	280
Total	**903**	**1091**	**915**
Number of fish measured:			
Yellowfin	33083	48858	46600
Skipjack	22121	23920	20573
Bigeye	7616	9307	8506
Bluefin	4922	6596	2567
Total	**67742**	**88681**	**78246**

TABLE 2. Comparison of the standard estimation values to those from species composition sampling. Values in **bold face** are significantly different.

	Standard catch	Species composition		
		Estimated	Low	High
2000				
Yellowfin	273245	268492	262305	274779
Skipjack	211252	**198449**	188412	208486
Bigeye	70153	**86755**	77849	95661
2001				
Yellowfin	396122	**388734**	381936	395532
Skipjack	145626	**136952**	129444	144460
Bigeye	42846	**58040**	51268	64812
2002				
Yellowfin	418967	**396763**	391480	402046
Skipjack	158043	153048	145489	160607
Bigeye	35201	**62396**	56334	68458