SOUTHWEST FISHERIES CENTER

August 1986

AN ANALYSIS OF BOTTOM FISH SIZE VARIATION AT THE HONOLULU FISH AUCTION

Stephen Ralston, Darryl T. Tagami, and Paul M. Shiota
Southwest Fisheries Center Honolulu Laboratory
National Marine Fisheries Service, NOAA
Honolulu, Hawaii 96822-2396

Not for Publication

ADMINISTRATIVE REPORT H-86-10

Southwest Fisheries Center Administrative Report H-86-10

AN ANALYSIS OF BOTTOM FISH SIZE VARIATION

AT THE HONOLULU FISH AUCTION

Stephen Ralston, Darryl T. Tagami, and Paul M. Shiota
Southwest Fisheries Center Honolulu Laboratory
National Marine Fisheries Service, NOAA
Honolulu, Hawaii 96822-2396

August 1986

NOT FOR PUBLICATION

INTRODUCTION

A fishery management plan (FMP) has been prepared by the Western Pacific Regional Fishery Management Council (WPRFMC) for the bottom fish and seamount fisheries of the region (Hawaii, American Samoa, Guam, and the Commonwealth of the Northern Marianas). This plan was prepared under the authority of the Magnuson Fishery Conservation and Management Act of 1976 and is intended to result in the management of these fisheries so that optimal yields are obtained. The FMP recently passed all phases of review and is now scheduled to be implemented and in effect by September 1986.

The plan stipulates that every year a monitoring team will assess the biological and economic conditions prevailing in these fisheries and an annual report will be prepared and presented to the WPRFMC based on its findings. A critical element of this report will be a section based on the analysis of bottom fish size-frequency data, the results of which will be used to update and revise our thinking concerning the extent of biological overfishing. An analysis of this type is specifically called for in the FMP and could, in principle, utilize the daily records of bottom fish sales at the United Fishing Agency (UFA) in Honolulu, Hawaii. For example, Ralston and Kawamoto (1985) analyzed the 1984 size structure of Hawaii's commercial landings of opakapaka, _Pristipomoides filamentosus_, using such data.

The purpose of this report is to examine the nature of size variation in parcels of bottom fish sold at the UFA to establish a stronger statistical basis for estimating the size structure of species in the bottom fish management unit. If representative size-frequency data can be obtained by examining individual UFA lot statistics, it will be possible to meet the reporting requirements of the bottom fish FMP without implementing a new and costly sampling program.

BOTTOM FISH TRANSACTIONS AT THE UNITED FISHING AGENCY

At the UFA a fisherman's bottom fish catch is auctioned off in either one of two forms, i.e., as individual fish or more commonly in "lots." A lot is composed of a grouping or collection of conspecific fish from a single fisherman's landings. Significantly, the fish are sorted so that all those within a single lot tend to be of similar size.

The fish in a lot are loaded either into metal tubs or onto rubberized black plastic pallets. Lots are identified with a slip of paper containing four pieces of information: (a) the species, (b) the total lot weight, (c) the number of individual fish, and (d) the fishing vessel landing the catch. Single fish lots are similarly tagged but for obvious reasons contain only items (a), (b), and (d). The fish are then placed on the auction floor for sale. After a buyer has bid and purchased a lot the company name and the bid price are also recorded. A typical bottom fish lot auctioned at the UFA contains an average of 6 fish and weighs approximately 18 kg (40 lb) (unpublished 1985 data).

Ralston and Kawamoto (1985) exploited the fact that fish are sorted by size to estimate the size structure of the 1984 landings of opakapaka at the UFA. They started with the premise that if fish were sorted perfectly by size there would be no variation in weight among individuals within a lot. The mean weight of fish within a lot would then be equal to the specific weights of all the individual fish which together comprised it. They showed that, in fact, there is size variation within lots, although it is small relative to overall levels of variation. For example, an analysis of variance of 694 individually weighed uku, Aprion virescens, taken from a sample of 167 UFA lots revealed that 91% of the total variation in weight was due to differences among lots, whereas only 9% of weight variation existed within lots. They argued that because variation in size of fish within lots is small relative to the total amount, as a first approximation one could regard all of the fish within a lot to be of equal size. Using this assumption the year's overall size composition of opakapaka landings was reconstructed by calculating the mean size of fish in each lot (total weight / total number of fish) and by accumulating an aggregate weight-frequency distribution by assigning each individual fish that passed through the auction to rounded 1-pound weight categories.

This type of approach rests heavily on the assumption that variation in fish weight within a lot is negligible. Were it possible to specify not only the mean weight of fish within a lot, but also the variance in weight and the form of the distribution, it would be a relatively straight forward calculation to estimate the overall size distribution of fish in each individual lot. Furthermore, based upon the information provided on auction lot sales slips, there are three factors which might be useful in developing a predictive equation to estimate variation in weight within a bottom fish lot: (a) the total lot weight, (b) the number of pieces, and (c) the species.

METHODS

A bottom fish subsampling program was instituted at the UFA beginning in December 1985 that continued through April 1986. The sampling was designed to gather data to develop a multiple regression equation for predicting weight variation within lots of species in the management unit, including the following:

Opakapaka	Pristipomoides filamentosus
Onaga	Etelis coruscans
Ehu	Etelis carbunculus
Uku	Aprion virescens
Hapuupuu	Epinephelus quernus
Butaguchi (pig ulua)	Pseudocaranx dentex

These six species in aggregate accounted for 73% of all lots and 84% of the weight of all bottom fish sold at the UFA in 1985 (unpublished data).

At the beginning of the subsampling program approval to sample on the premises was granted by the management of the UFA. Individual buyers and wholesalers were approached and asked to participate in the project by allowing their fish to be individually weighed. By the end of the subsampling effort a number of wholesalers had cooperated, including the following companies: Wing Sing, Star, Tropic, Drum, HSP, J & F, Fishland, Maeda, Nishimura, Fresh Exports, KK, and Red & White.

After identifying lots purchased by one of the participating buyers, all of the fish that comprised it were weighed and the basic lot sales slip statistics were recorded (see above). These data, including the transaction date, were entered into standardized data collection sheets, which were subsequently coded, entered into the Honolulu Laboratory computer, and verified. The data were analyzed with Statistical Analysis System (1979) computer routines and in particular with PROC GLM, a procedure that uses the principle of least squares to fit generalized linear models. It performs univariate and multivariate analyses, including simple linear regression, multiple linear regression, analysis of variance, analysis of covariance, and partial correlation analysis.

Although it is preferable to use the metric system for fishery sampling programs, the UFA records the weights of bottom fish in pounds. Furthermore, fishermen are accustomed to discussing the size of fish they catch in pounds and their understanding of size distributions is also based on this measure. Lastly, current State of Hawaii law governing the take of undersized bottom fish is expressed in pounds. For these reasons the basic unit of weight used in this study was the pound measured to the nearest ounce, although the metric equivalent in kilograms is provided when appropriate.

RESULTS

A total of 478 lots which in aggregate were composed of 3,159 fish were measured, including the uku that had been weighed previously (Ralston and Kawamoto 1985). The breakdown by individual species was as follows:

Species	Lots	Fish
Opakapaka	77	809
Onaga	116	772
Ehu	36	220
Uku	153	838
Hapuupuu	37	186
Butaguchi	59	334

These data show that an average of 80 lots comprising 526 fish were measured for each of the 6 species of interest.

The data were first examined to determine the form of the underlying weight distribution within bottom fish lots. For each lot the mean ($\bar{X}$) and standard deviation (s) were calculated from the weights of fish within it. Individual fish weights within a lot (X) were then converted to standard scores according to the following formula:

$$Z = \frac{X - \bar{X}}{s}$$

Z-scores were rounded to the nearest 0.25 and tallied into a pooled frequency distribution because preliminary inspection of the data on a species by species basis revealed no differences. The aggregate histogram of all the data is presented in Figure 1. It should be noted that lots containing exactly two fish were not included in this analysis. This is because after calculating the mean and standard deviation from any sample where N = 2, the untransformed values will always be converted into Z-scores equal to -0.71 and +0.71. The effect is to artificially increase the frequency of standard scores in the intervals that encompass these values, giving the superficial appearance of a bimodal distribution where none exists.

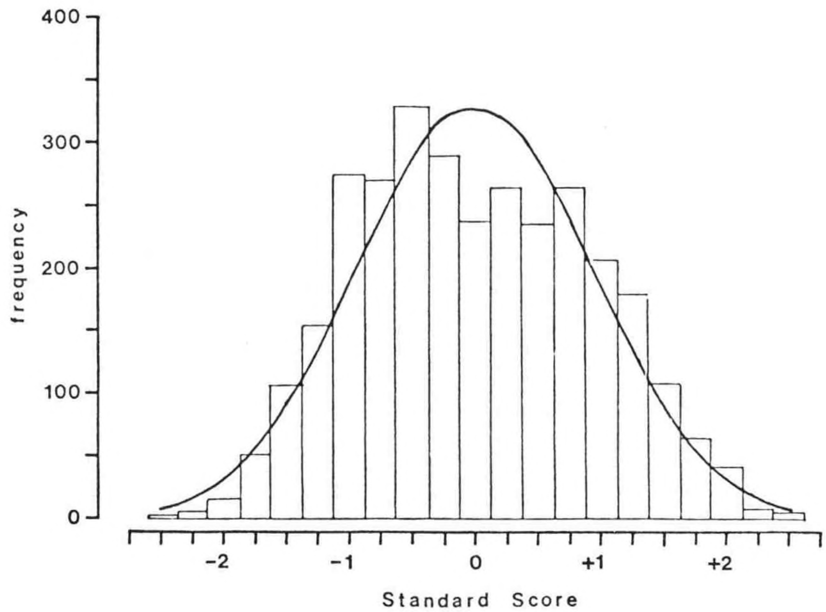

Figure 1. Observed and theoretical frequency distributions of Z-scores derived from within lot subsampling.

In the figure the observed distribution of Z-scores (N = 3109) is presented as a histogram. Superimposed on this is a normal distribution, with mean and standard deviation equal to the sample statistics calculated from the observed frequency data ($\bar{Z} = 0.00$, $s = 0.92$). Although the observed distribution of standard scores departs significantly from the normal curve (Kolmogorov-Smirnov D statistic = 0.047, $P < 0.01$) it is evident that the actual underlying distribution of weights within bottom fish lots is symmetrical and is near normal. Calculation of the third and fourth moments about the mean indicate that the distribution shows slight positive skewness ($a_3 = 0.150$) and is platykurtic ($a_4 = -0.728$). Nonetheless, the normal distribution represents an adequate approximation to the data, which could be used as a theoretical probability model for describing weight variation within lots of bottom fish.

Ralston and Kawamoto (1985) reported that 91% of the variation in weight of uku sold at UFA was attributable to differences among lots. To similarly determine the balance between "within" and "among" lot weight variation in this study, standard analysis of variance (ANOVA) was applied to the data. A separate ANOVA was run for each species as well as for all the data combined. In all analyses the bottom fish lot was the treatment variable. For each ANOVA the coefficient of determination was:

Species	r^2
Opakapaka	90.5%
Onaga	96.3%
Ehu	96.5%
Uku	89.3%
Hapuupuu	93.6%
Butaguchi	88.0%
All species pooled	94.3%

These findings show clearly that the preponderance of variation in bottom fish weight is attributable to differences among lots. When the data were pooled, for example, 94.3% of total weight variation was "explained" by differences in the mean weight of fish sampled from each lot. These data therefore support the assumption of Ralston and Kawamoto (1985) that, as a first approximation, within lot variation can be considered negligible.

The data were examined further to develop a suitable regression model for estimating variance in weight given the following basic lot characteristics: (a) the total number of fish (#FISH), (b) the total weight of fish in the lot (TOTWT), and (c) the species. To accomplish this the variance of weight within a lot of fish (VAR) was calculated for each of the 478 lots. Scattergrams were then prepared of VAR on #FISH and TOTWT. These showed curvilinearity in the relationship of these variables. Natural logarithms of the data were determined. Plots resulting from the transformed data showed that the relationship of log(VAR) to log(#FISH) and log(TOTWT) was linear (Figs. 2 and 3). The following regression model was subsequently developed for further treatment of the data:

$$(1) \quad \log(\mathrm{VAR}) = a \cdot \log(\#\mathrm{FISH}) + b \cdot \log(\mathrm{TOTWT}) + \mathrm{Intercept} + \mathrm{Normal\ Error}$$

which is equivalent to a multiplicative model of the form:

$$\text{Expected Value of VAR} = (C_o)(\#\mathrm{FISH}^a)(\mathrm{TOTWT}^b)$$

where C_o is a constant, the antilog of the intercept from the log-log regression.

Equation (1) is specified as a multiple regression equation in which log(#FISH) and log(TOTWT) are the independent variables. To assess the orthogonality (i.e., independence) of these variables species-specific scattergrams and correlations were developed. The results showed that for the onaga, ehu, and hapuupuu the variables are independent, whereas for the

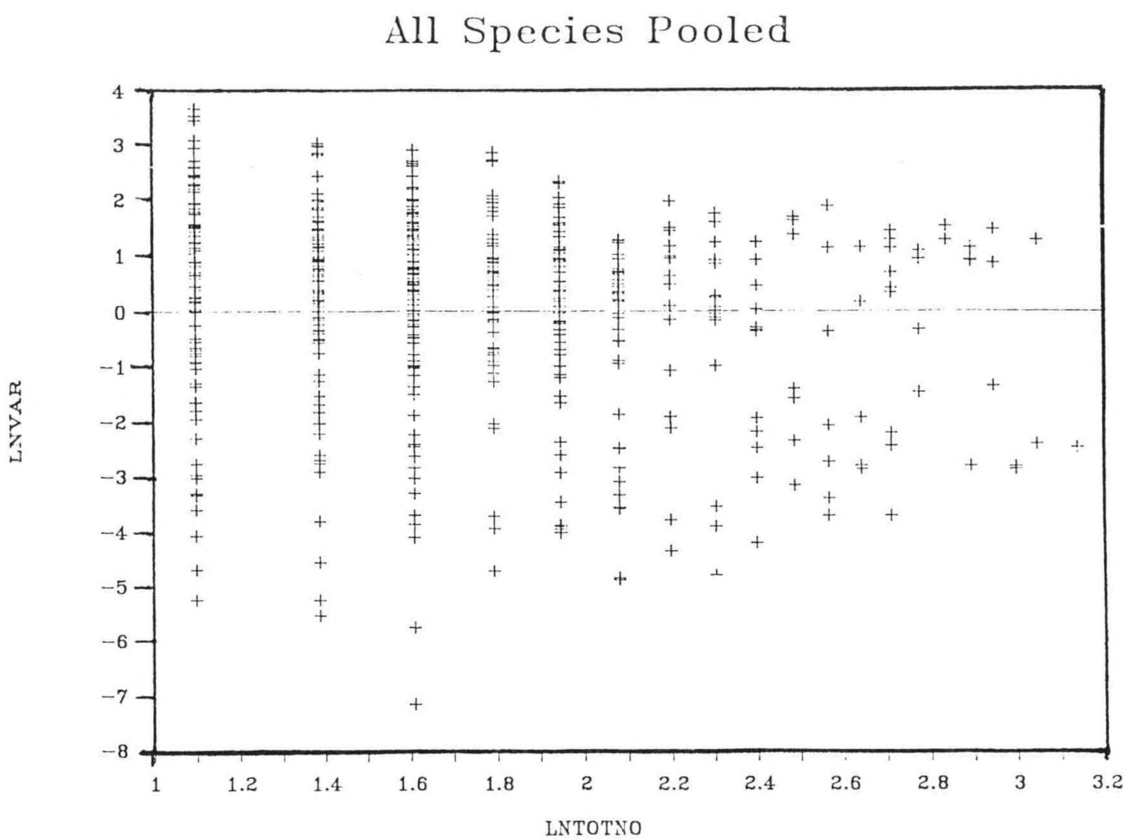

Figure 2. Scattergram of log(VAR) against log(#FISH).

All Species Pooled

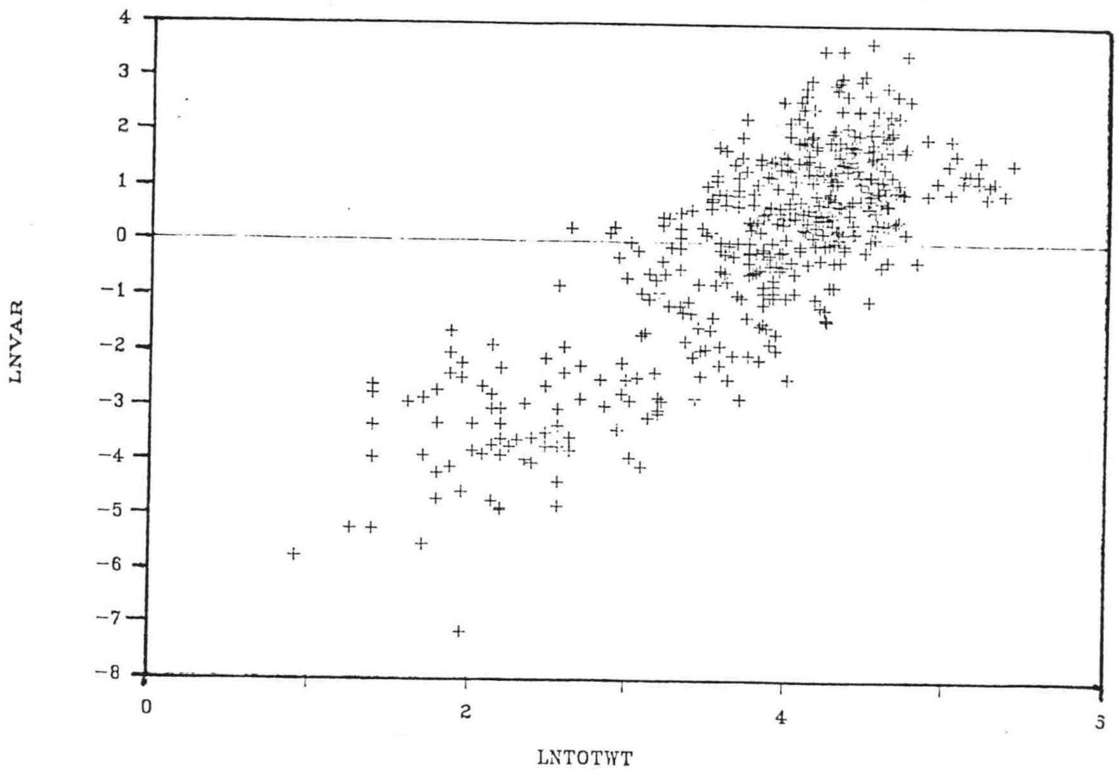

Figure 3. Scattergram of log(VAR) against log(TOTWT).

opakapaka (r = 0.69), uku (r = 0.77), and butaguchi (r = 0.46) significant positive correlations exist in the relationship of log(#FISH) and log(TOTWT). These findings indicate that caution must be exhibited in the use of multiple regressions involving the latter three species, to insure that the independent variables remain within the range of data for which the model was initially fitted.

The data were next analyzed by analysis of covariance (ANCOVA) to determine whether differences exist among bottom fish species with regard to the multivariate relationship specified in Equation (1). A similar relationship among species might be expected solely on the basis of similarities in body size and shape, this because weight variation within lots is largely the result of how well fish have been sorted by UFA staff. Therefore, in this analysis the six species were used to specify the level of the classification or treatment variable. Log(#FISH) and log(TOTWT) were employed as covariates and log(VAR) was the dependent variable. The ANCOVA showed that significant differences do exist among the species in the relationship of log(VAR) to the covariates ($F = 246.45$, df = 7, 445, $P < 0.0001$, multiple $r^2 = 0.79$). The treatment means, adjusted for initial differences in levels of the covariates and with standard errors, are:

Species	Adjusted mean	Standard error
Hapuupuu	0.4251	0.1476
Ehu	0.1963	0.1690
Uku	0.0094	0.0791
Butaguchi	-0.0056	0.1174
Opakapaka	-0.1189	0.1088
Onaga	-0.4282	0.0810

These results show that when species are compared at equivalent values of #FISH and TOTWT, weight variation within lots is greatest for the hapuupuu. Onaga lots appear to contain the least amount of variation.

Comparison of the adjusted means was used to determine which species are similar to one another and which are dissimilar. The following matrix presents the probabilities of greater t-values under the null hypotheses that the adjusted means for species i (rows) and species j (columns) are equal. Although these multiple comparisons result in an increased chance of a Type I error (rejection of a true null hypothesis), the consequences of Type I error are acceptable-whereas for Type II errors they are not.

	SPECIES					
SPECIES	Hapuupuu	Ehu	Uku	Butaguchi	Opakapaka	Onaga
Hapuupuu	–					
Ehu	0.3312	–				
Uku	0.0114	0.3296	–			
Butaguchi	0.0190	0.3542	0.9135	–		
Opakapaka	0.0043	0.1116	0.3597	0.4897	–	
Onaga	0.0001	0.0008	0.0001	0.0034	0.0234	–

The probabilities given above indicate no significant differences ($P = 0.05$) in the adjusted means of butaguchi, ehu, opakapaka, and uku. However, onaga is distinguishable from all other species and hapuupuu is different from all other species except ehu. This result suggests the following species partition for calculating three distinct multiple regression equations for estimating variation in weight of bottom fish within auction lots at UFA:

Equation A:	Opakapaka, ehu, uku, butaguchi
Equation B:	Onaga
Equation C:	Hapuupuu

All species listed for Equation A have similar adjusted mean log(VAR), i.e., no member of the group is different from any other member. This result was confirmed by a second ANCOVA involving the data for only these four species. Again, in no case did any one species differ from another.

The data were then aggregated according to the groupings specified above (A, B, and C) and the multiple regression coefficients of Equation (1) estimated for each. All three regressions were highly significant ($P < 0.0001$) and each independent variable (log(#FISH) and log(TOTWT)) accounted for a significant reduction in the error sums of squares in all three cases. Parameter estimates with associated standard error estimates given in parentheses below are:

Species	a	b	Intercept	r^2	N
Opakapaka, ehu, Uku, butaguchi	−1.4402 (0.1026)	1.9906 (0.0594)	−5.0473 (0.2363)	79%	300
Onaga	−1.5431 (0.1948)	2.0978 (0.1118)	−5.6953 (0.5573)	79%	116
Hapuupuu	−1.2659 (0.3880)	1.4321 (0.2475)	−2.5497 (1.2068)	57%	37

The r^2 values for the first two regressions indicate that roughly 80% of the variation in weight of these species within UFA bottom fish lots is "explained" by the regressions. This figure, coupled with the r^2 values estimated earlier by ANOVA for among lot variation (94%), suggests that nearly 99% of the total variation in weight of bottom fish can be accounted for with the present model. Similarly, for hapuupuu approximately 97-98% of its total weight variation can be recovered from simple lot statistics (i.e., the total number of fish, total lot weight, and the species).

EXAMPLE APPLICATION OF THE METHOD

Following is an application of these results to estimate the distribution of weights within a lot of UFA bottom fish. Suppose that a lot of opakapaka weighs 47 lb (21.3 kg) and comprises 8 fish. From these data we would estimate the mean weight of fish within this lot to be 5.88

lb (2.67 kg). An estimate of the variance in weight could be calculated as follows:

$$\log(VAR) = -1.4402 \log(\#FISH) + 1.9906 \log(TOTWT)) -5.0473$$

$$\log(VAR) = -1.4402 \log(8) + 1.9906 \log (47) - 5.0473$$

$$\log(VAR) = -1.4402 (2.08) + 1.9906 (3.85) - 5.0473$$

$$\log(VAR) = -2.99 + 7.66 - 5.0473$$

$$\log(VAR) = -0.38$$

$$VAR = 0.69$$

The standard deviation of weights within the lot is s = sqrt(VAR) = 0.83 lb (0.38 kg). Given that the distribution is normal (see Figure 1) we can estimate the probabilities that fish are smaller than any specific weight by using standard scores as follows:

Weight (lb)	Z Score	Probability $\leq$ Z
3.5	-2.87	0.0021
4.5	-1.66	0.0485
5.5	-0.46	0.3228
6.5	+0.75	0.7734
7.5	+1.95	0.9744
8.5	+3.16	0.9992

Thus the probability that a fish lies in the interval 3.5-4.5 lb, i.e., that its rounded weight is 4 lb, is 0.0485 - 0.0021 = 0.0464. Similarly, the remaining probabilites are calculated as:

Weight	Probability	Expected Number
4	0.0464	0.37
5	0.2743	2.19
6	0.4506	3.60
7	0.2010	1.61
8	0.0248	0.20
	Sum =	7.97 $\simeq$ 8.00

The number of fish expected at a given weight is simply the probability of occurrence at that weight times the total number of fish in the sample (8). The last column therefore gives the expected number of fish in each weight category.

This procedure can be employed to estimate the full weight-frequency distribution of all six species for which auction data is supplied. These estimated size distributions can then be employed in stock-assessment models *as if* the entire market sample had been measured. The error in doing so should be acceptably small.

Acknowledgments

This study would never have been possible without the full cooperation of the United Fishing Agency and the numerous buyers and wholesalers of fresh fish who attend the auction every day. We are especially indebted to Brooks Takanaka for his patience in coordinating our efforts while managing activities on the auction floor and for his help in assuring the study reached a successful conclusion.

Literature Cited

Ralston, S., and K. Kawamoto.
1985. A preliminary analysis of the 1984 size structure of Hawaii's commercial opakapaka landings and a consideration of age at entry and yield per recruit. NOAA, NMFS, SWFC, Honolulu Laboratory Admin. Rep. H-85-1, 9 p.

Statistical Analysis System.
1979. SAS user's guide. SAS Institute Inc., Raleigh, N. C., 494 p.