SAMPLE SIZE FOR ESTIMATING DOLPHIN MORTALITY ASSOCIATED WITH THE TUNA FISHERY

NANCY C. H. LO, National Marine Fisheries Service, P.O. Box 271, La Jolla, CA 92038

Abstract: Dolphin (*Stenella* spp., *Delphinus delphis*) mortality caused by U.S. tuna purse seiners in the eastern tropical Pacific has been monitored under a quota system since 1976. A stratified ratio estimator (kill-per-day) has been used to measure the kill rate during the year. Vessel trips are stratified by fishing time period. The mortality data from 1976 through 1978 suggest a sample-size procedure based on the coefficient of variation rather than on the absolute variance of kill data. The sample size for each stratum is computed according to the desired precision of the stratum estimates rather than by allocation of a yearly sample size. The probability that the estimated mortality exceeds the quota, when in fact the quota is higher than the true mortality, is computed. This sample-size procedure is recommended for data that have a standard deviation proportional to the mean and the situation where the estimate within stratum is as important as the overall estimate.

J. WILDL. MANAGE. 47(2):413–421

Yellowfin tuna (*Thunnus albacares*) are often associated with dolphins (e.g., *Stenella attenuata*, *S. longirostris*, etc.) in the eastern tropical Pacific (ETP) and fishermen have used this association to catch fish. A purse seine net (McNeely, Pac. Fish. 59[7]:27–58, 1961) is used to encircle an entire aggregation of tuna and dolphin in an operation called a "dolphin set." During the dolphin set, dolphins encircled together with tuna are often entangled in the nets and accidentally drown.

The purse seine fishing technique in the ETP was first introduced to the tuna fleet in the early 1960's. Currently there are about 100 purse seiners, each with a carrying capacity of over 400 short tons, in the U.S. fleet. Each vessel averages 3 or 4 trips a year and the entire fleet is expected to make about 350 trips resulting in 6,000 dolphin sets each year.

The fishing area is from 40°S to 40°N latitude, 150°W longitude to the west coast of the American continent. The yellowfin tuna catch in the area known as the Commission Yellowfin Regulatory Area (CYRA) is regulated by an allowable tuna catch set by the Inter-American Tropical Tuna Commission (IATTC). After the allowable tonnage catch of yellowfin tuna has been taken (usually by Jul) in the CYRA, most vessels fish outside the CYRA and mainly target on tuna associated with dolphins. Hence, higher dolphin mortality is expected after July.

The National Marine Fisheries Service (NMFS) initiated an observer program in 1971 through which NMFS-trained biological technicians are placed aboard randomly selected U.S. tuna vessel trips to collect data on dolphin mortality and biological information. Kill rates of the observed and unobserved vessels are assumed to be similar because the bias, if it exists, cannot be measured.

Until 1977, it was estimated that over 100,000 dolphins were killed each year. In 1976, the NMFS established an annual quota on maximum allowable kill of dolphins by U.S. fishermen. The objective of this paper is to discuss a sample-size procedure for a stratified-ratio estimator of the incidental dolphin mortality on a real-time basis during the year with vessel trips stratified by fishing time period. This procedure accounts for decreasing mortality levels in recent years resulting from gear improvement, as well as the high associ-

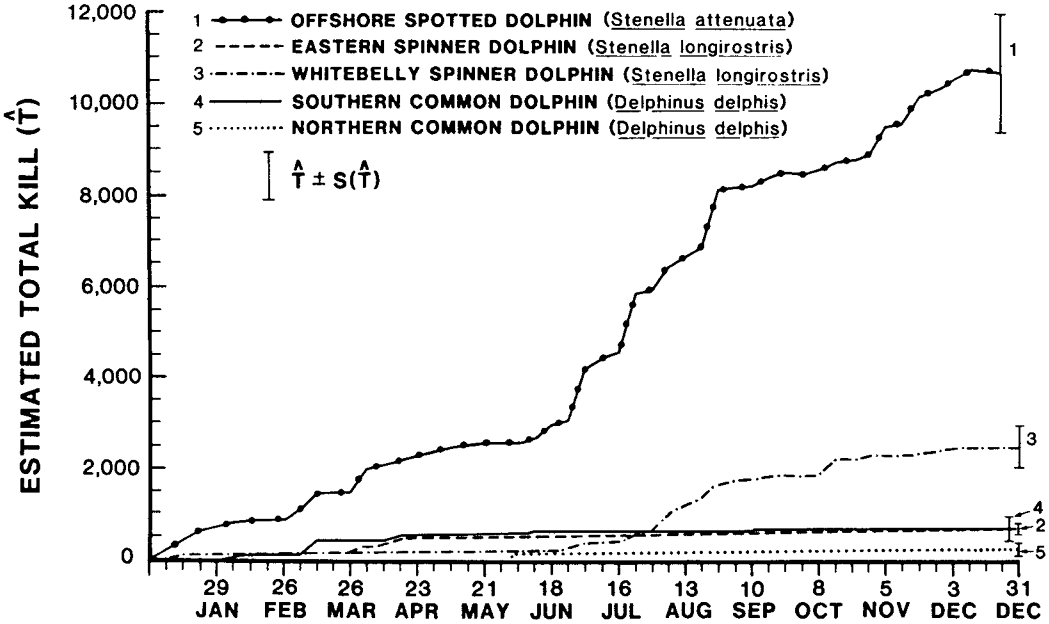

Fig. 1. Weekly estimates of cumulative mortality by management stock of dolphins, 1978.

ation of mean and standard deviation of the kill. This procedure is based on the coefficient of variation (CV) of the estimators of kill rates rather than on their absolute variances (Cochran, Sampling techniques, J. Wiley & Sons, New York, N.Y., 1977).

Two anonymous referees and C. E. Braun made valuable suggestions for improvement of the manuscript. Their help is appreciated.

MORTALITY MONITORING PROGRAM

The Marine Mammal Protection Act of 1972 and a court ruling in 1976 required the NMFS to set an annual quota for the incidental take of dolphins by the American tuna fleet. The quota is an attempt to reduce incidental dolphin mortality associated with tuna fishing by U.S. purse seiners in the ETP. For 1976–80, the annual quotas were 78,000, 62,000, 52,000, 42,000, and 31,000, respectively. After 1976, the quotas were set by species (stock). If the quota of any species (stock) is reached, the yellowfin tuna fishing on that species (stock) is prohibited for the remainder of the year (U.S. Dep. Commer., Fed. Reg. Pt. 3, 42[40]:12010–12020, 1977).

Beginning in 1976, biological technicians aboard tuna vessels reported cumulative mortality data each week by radio. The observed kill-per-day is used to estimate the total mortality for the entire fleet on a real-time basis (Fig. 1). Many kill ratios have been used to measure kill levels, e.g., kill-per-set, kill-per-tonnage-of-yellowfin tuna-catch. Kill-per-day was chosen to estimate total kill during the year because data on total days are available throughout the year whereas tonnage of yellowfin tuna catch and total number of dolphin sets are not known until the end of the year (Lo et al., Fish. Bull. 80:396–400, 1982).

DOLPHIN MORTALITY ESTIMATION

The simplest estimate of total mortality to a given point in time would be the total mortality for the observed vessels inflated by the ratio: number of vessels in the fleet divided by number of observed vessels. However, because each vessel does not make the same number of fishing trips and trips vary in length, kill-per-trip or kill-per-day may be used to reduce the variability in the estimate of total mortality. The kill-per-day statistics have been used for monitoring the mortality against the quota. In a stratified sampling scheme, the estimated total mortality based on kill-per-day can be derived as described below.

The sampling units are vessel trips stratified by fishing time period to account for both average kill and the availability of kill differing among seasons within a year. The season is determined by the availability of yellowfin tuna and the allowable catch tonnage for yellowfin tuna set by the IATTC.

A year is divided into 3 periods (Jan–Mar, Apr–Jun, Jul–Dec). The number of trips starting within each period is considered as the total number of vessel trips within stratum (N_i, $i = 1, 2, 3$). The trip data from previous years are used to predict N_i for the current year. The vessel trip distribution (N_i) with percentage in parentheses in 1978 was

$i = 1$	2	3	
(Jan–Mar)	(Apr–Jun)	(Jul–Dec)	Total
115 (35)	94 (28)	122 (37)	331 (100).

Note that for the ith time period:

n_i = number of observed trips,
x_{ij} = number of dolphins killed by the jth observed (sampled) vessel trip,
d_{ij} = number of days at sea (duration) of the jth observed vessel trip,
N_i = total number of vessel trips,
f_i = n_i/N_i = sampling intensity,
D_i = total number of days at sea by the fleet,
T_i = total kill,
r_i = kill-per-day = $T_i/D_i = \mu_i(x)/\mu_i(d)$, where $\mu(\cdot)$ is the true mean, i.e., $\mu_i(x) = T_i/N_i$ and $\mu_i(d) = D_i/N_i$, and
T_i = $D_i r_i$.

Total kill for the year is the sum of total kill within each time period, i.e.,

$$T = \sum_i T_i = \sum_i D_i r_i$$

and

$\hat{T}_i$ = the estimated total kill in the ith time period = $D_i\hat{r}_i$.

To compute $\hat{T}_i$, it remains to choose appropriate estimates $\hat{r}_i$ for r_i, the "average kill-per-day" in the ith stratum. The data obtained from the observed vessels are number of kills x_{ij} by the jth vessel trip in the ith stratum. It is assumed that r_i is related to the kill x_{ij} by

$$x_{ij} = r_i d_{ij} + e_{ij},$$

where e_{ij} has mean zero and variance σ_i^2:

$$\sigma_i^2 = \sum_{j=1}^{N_i} (x_{ij} - r_i d_{ij})^2/N_i.$$

The estimator of the form

$$\hat{r}_i = \sum_j x_{ij}/\sum_j d_{ij} = \bar{x}_i/\bar{d}_i$$

is used because $\hat{r}_i$ is a consistent estimator of r_i in a sense that

$$\lim_{n_i \to N_i} \hat{r}_i = r_i.$$

Moreover, 1977 data indicated a linear relationship between kill (x_{ij}) and days at sea (d_{ij}) (Fig. 2), and the variance of kill around the line is proportional to number of days at sea (d_{ij}). For this model

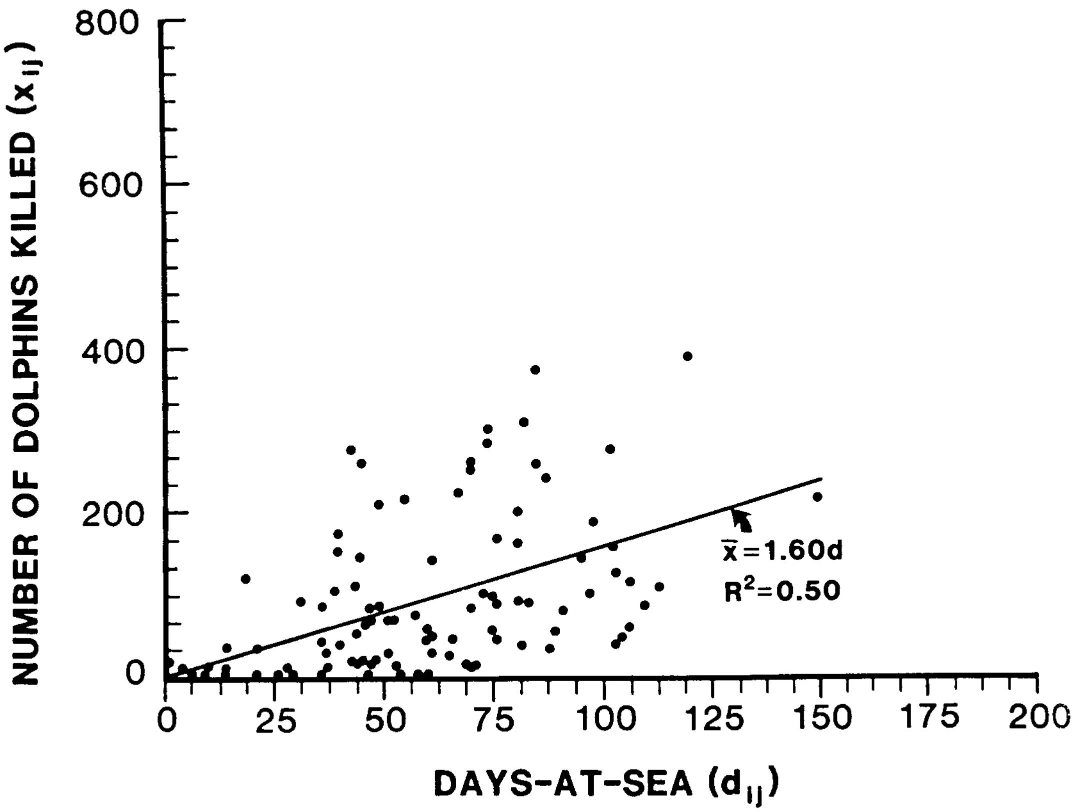

Fig. 2. Linear relationship of dolphin kill and days at sea, 1977.

$$\hat{T} = \sum_i D_i \hat{r}_i = \sum_i D_i \sum_j x_{ij} / \sum_j d_{ij}. \quad (1)$$

The variance of $\hat{r}_i$ according to Cochran (1977) is

$$\sigma^2(\hat{r}_i) = (1 - f_i)\frac{1}{n_i}\frac{1}{\mu_i^2(d)}\sigma_i^2,$$

where σ_i^2 is estimated by

$$s_i^2 = \sum_{j=1}^{n_i} (x_{ij} - \hat{r}_i d_{ij})^2/(n_i - 1). \quad (2)$$

Thus $\sigma^2(\hat{r}_i)$ is estimated by

$$s^2(\hat{r}_i) = (1 - f_i)\frac{1}{n_i}\frac{1}{\mu_i^2(d)}s_i^2 \quad (3)$$

and the sample variance of $\hat{T}$ is, assuming that $\hat{T}_i$ and $\hat{T}_j$ are independently distributed:

$$s^2(\hat{T}) = \sum_i D_i^2 s^2(\hat{r}_i). \quad (4)$$

SAMPLE SIZE DETERMINATION

Determination of sample size for a stratified sampling scheme could be based on the magnitude of both the mean and variance of kill of each stratum obtained in the previous years. However, this procedure is applicable only when the within-stratum variability of kill is independent of the average kill-per-trip and thus may not be valid for dolphin mortality estimation.

One characteristic of dolphin mortality data was the high correlation between mean kill and standard deviation. Because the mean kill varies for vessels with different gear type, carrying capacities, and fishing time periods, the observed vessels in 1976–78 were grouped according to these categories. To illustrate the high correlation between mean and standard deviation of dolphin mortality, average kill-per-trip ($\bar{x}$) and standard deviation ($\bar{s}$) were computed for each group and fitted by simple linear regression through the origin: $s = 1.37\bar{x}$ and $R^2 = 0.93$. Note that the regression coefficient 1.37 ($=s/\bar{x}$) was also an estimate of the coefficient of variation of the kill data ($CV = \sigma/\mu$) over a 3-year period.

Kill rates have dropped each year as a result of the dolphin mortality reduction effort; the unstratified kill-per-trip estimates for 1976, 1977, and 1978 were approximately 200, 106, and 63, respectively. The corresponding standard deviations of the kill were 290, 120, and 124. Thus, the estimate of individual yearly coefficient of variation was 1.45, 1.13, and 1.97, respectively. To obtain the same level of precision in each year (e.g., 95% CI = estimate $\pm\ 2 \times 0.10 \times$ estimate), more observed trips would be required for 1978 than for 1977 even though the absolute variances for 1978 and 1977 were similar. The fact that yearly distributions of the observed mortality are skewed and that the standard deviation is usually larger than the mean necessitates large sample sizes.

A stratified sampling scheme usually attempts to maximize the precision of the estimate of the population total over all strata. For dolphin mortality, however, the mortality estimates in the 1st 2 periods are also important because they must be compared directly with the yearly quota. Therefore, the number of observed trips should be determined within each stratum rather than by the standard method of determining n, total number of observed trips, which is allocated to each stratum (Cochran 1977).

An individual stratum is each fishing period within a year (e.g., Jan–Mar). Within each stratum, the number of observed trips (sample size) was calculated. For stratum i, the sample size required for a desired $CV(\hat{T}_i)$ was written in terms of the estimated $CV(x_{ij})$ available from historical data: $CV_i = s_i/\bar{x}_i$ and s_i is the sample standard deviation of kill computed according to eq. (2).

Using eqs. (1) and (3) (Cochran 1977),

$$\begin{aligned} CV(\hat{T}_i) &= \frac{\sigma(\hat{T}_i)}{\mu(\hat{T}_i)} \doteq \frac{s(\hat{T}_i)}{\hat{T}_i} = \frac{s(D_i\hat{r}_i)}{D_i\hat{r}_i} = \frac{s(\hat{r}_i)}{\hat{r}_i} \\ &= \left[\frac{1}{n_i}(1-f_i)\right]^{\frac{1}{2}} \frac{s_i}{\mu_i(d)} \frac{1}{\hat{r}_i} \\ &\doteq \left[\frac{1}{n_i}(1-f_i)\right]^{\frac{1}{2}} \frac{s_i}{\bar{x}_i} \\ &= \left[\frac{1}{n_i}(1-f_i)\right]^{\frac{1}{2}} CV_i. \end{aligned} \tag{5}$$

Solving eq. (5) for n_i,

$$\begin{aligned} n_i &= \left\{\frac{1}{N_i} + \left[\frac{CV(\hat{T}_i)}{CV_i}\right]^2\right\}^{-1} \\ &= \left(\frac{1}{N_i} + \phi_i^2\right)^{-1}, \end{aligned} \tag{6}$$

where $\phi_i = [CV(\hat{T}_i)]/CV_i$. $CV(\hat{T}_i)$ is specified by the user and CV_i is estimated from past data. The sample size (n_i) is independent of the absolute value of mean and variance. For dolphin mortality estimation, the mean and variance of kill change each year. Equation (6) is used to compute the sample sizes.

Sample size for the entire year is the sum of n_i ($i = 1, 2, 3$). The number of trips per stratum as estimated from eq. (6) depends on total number of trips (N_i), and the ratio (ϕ_i) of $CV(\hat{T}_i)$ and CV_i. The min-

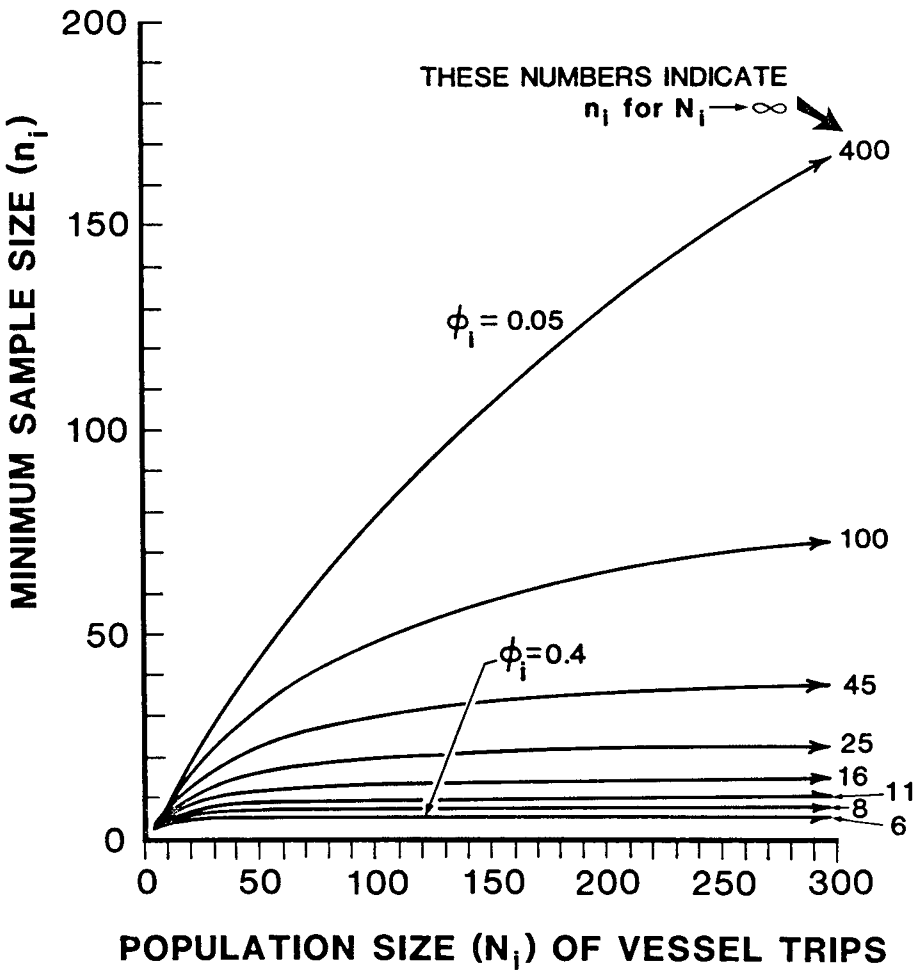

Fig. 3. Minimum sample sizes for population sizes (N_i) and ϕ_i, ϕ_i = 0.05 (0.05) (0.40).

imum number of observed trips was computed for N_i up to 350 and ϕ_i = 0.05 (0.05) (0.4) (Fig. 3).

PROBABILITY THAT ESTIMATED MORTALITY EXCEEDS ESTABLISHED QUOTA

The annual quota is set for each species (stock). For any 1 species (stock), if the estimated mortality ($\hat{T}$) exceeds its quota (Q), i.e., $\hat{T} > Q$, then the yellowfin tuna fishing associated with that species (stock) is prohibited for the remainder of the year. Sample sizes for monitoring purposes were evaluated based on the probability that the estimated total mortality exceeds or falls short of an established quota when in fact the true total mortality is below or above the quota. If the true total mortality is at or slightly below quota levels, overestimated mortality is likely but undesirable. When the true total mortality is much below the quota, observed trips may be reduced because only crude estimates of mortality would be required. This probability of risk depends on the sampling distribution of the estimated total mortality, which is a function of sample size and the variation of the data, measured by coefficient of variation. Under a stratified sampling scheme the probability of an erroneous decision (P) is denoted by:

$$P[\hat{T} > Q \mid T < Q, n_i, \mathrm{CV}_i, \mathrm{CV}(\hat{T}_i)] \text{ or}$$
$$P[\hat{T} < Q \mid T > Q, n_i, \mathrm{CV}_i, \mathrm{CV}(\hat{T}_i)],$$

where T = actual total mortality by all vessel trips,
$\hat{T}$ = estimated total mortality,
Q = quota on allowable total mortality of dolphins,
n_i = total number of observed trips within stratum i, and
CV_i and $\mathrm{CV}(T_i)$ are defined as before.

Assuming the distribution of total mortality during the year is proportional to the yellowfin catch on dolphin over time, it is clear that $T = T(g_1 + g_2 + g_3) = T_1 + T_2 + T_3$, where g_i is the proportion of yellowfin catch on dolphin in the ith time period during the year and is predicted based on past data, $\sum_i g_i = 1$.

Likewise, $\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3$

and

$$\sigma^2(\hat{T}) = \Sigma\, \sigma^2(\hat{T}_i) = \Sigma\, T_i^2\mathrm{CV}^2(\hat{T}_i)$$
$$= T^2\, \Sigma\, g_i^2\mathrm{CV}^2(\hat{T}_i).$$

If it is assumed that $\hat{T}$ is normally distributed with mean T and variance $\sigma^2(\hat{T})$ estimated by $s^2(\hat{T})$, then

$$P[\hat{T} > Q \mid T < Q, n_i, \mathrm{CV}_i, \mathrm{CV}(\hat{T}_i)] \text{ or}$$
$$P[\hat{T} < Q \mid T > Q, n_i, \mathrm{CV}_i, \mathrm{CV}(\hat{T}_i)]$$
$$= 1 - \Phi\left[\frac{|Q - T|}{\sigma(\hat{T})}\right]$$
$$= 1 - \Phi\left\{\frac{|Q - T|}{T[\Sigma\, g_i^2\mathrm{CV}^2(\hat{T}_i)]^{\frac{1}{2}}}\right\}, \qquad (7)$$

where $\Phi(y)$ is the probability that a standard normal random variable exceeds y. Using eq. (5)

$$\mathrm{CV}(\hat{T}) = \frac{\sigma(\hat{T})}{T} \doteq \frac{T[\Sigma\, \mathrm{CV}^2(\hat{T}_i)g_i^2]^{\frac{1}{2}}}{T}$$
$$= \Sigma\, [\mathrm{CV}^2(\hat{T}_i)g_i^2]^{\frac{1}{2}}$$
$$= h[\Sigma\, g_i^2]^{\frac{1}{2}} \qquad (8)$$

if we choose $\mathrm{CV}(\hat{T}_i) = h$ for all i.

Thus, for a desired $\mathrm{CV}(\hat{T}_i) = h$ or n_i and given values of T, CV_i, g_i, and N_i, the probability P in eq. (7) can be computed. $\mathrm{CV}(\hat{T})$ may be obtained by eq. (8).

RESULTS

The sample size n_i for 1979 vessels was computed to illustrate the technique. The 1978 trip data were used to predict 1979 vessel trips: the trips per vessel and the trip distribution during the year, e.g., the number of trips starting within each period (N_i, i = 1 [Jan–Mar], 2 [Apr–Jun], and 3 [Jul–Dec]).

At the end of 1978, 86 vessels were anticipated to be fishing in 1979. The average number of trips per vessel is 3.18 and the estimated trip distribution for the 3 periods is 35, 28, and 37%, respectively. Therefore,

$$\begin{bmatrix} N_1 \\ N_2 \\ N_3 \end{bmatrix} = 86 \times 3.18 \times \begin{bmatrix} 0.35 \\ 0.28 \\ 0.37 \end{bmatrix} = \begin{bmatrix} 95 \\ 77 \\ 101 \end{bmatrix}.$$

Reasonable values of $\mathrm{CV}(\hat{T}_i)$ to consider are 0.10, 0.15, 0.20, and 0.25. CV_i is difficult to predict during any particular time period. However, using the range of CV_i observed in the past as a guideline, CV_i between 1.0 and 2.0 are probably representative. Thus, observed trips are allocated for 1979 (Table 1). A total of 273 trips was predicted for 1979; the minimum number of observed trips ranged from 41 to 223 depending on the required degree of precision of the estimates and the assumed variation of the data.

This sampling scheme first determines the sample size within stratum, and the

Table 1. Minimum number of observed trips (n_i) needed for each period to obtain mortality estimated within stratum with levels of precision (CV[$\hat{T}_i$] = 0.10, 0.15, 0.20, and 0.25) based on data with 3 levels of variation (CV_i = 1.00, 1.50, and 2.00). (Equation 6).

Time period (i)	N_i	CV($\hat{T}_i$) 0.10			0.15			0.20			0.25		
CV_i		1.00	1.50	2.00	1.00	1.50	2.00	1.00	1.50	2.00	1.00	1.50	2.00
ϕ_i		0.10	0.07	0.05	0.15	0.10	0.08	0.20	0.13	0.10	0.25	0.17	0.13
1	95	49	65	77	30	49	59	20	36	49	14	26	36
2	77	44	56	65	28	44	52	19	33	44	13	25	33
3	101	50	68	81	31	50	61	20	37	50	14	27	37
Totals	273	143	189	223	89	143	172	59	106	143	41	78	106

yearly sample size is the sum of stratum sample sizes (Table 1). Minimum sample sizes for other combinations of N_i and ϕ_i can be obtained from eq. (6) or Fig. 3.

The probability that $\hat{T}$ exceeds Q was calculated for the case where $\hat{T} < Q$ because estimates of historical kill rates indicated that mortality was below the 1979 quota (Table 2). Based upon the 1978 yellowfin catch on dolphin distribution over time (i.e., $g_1 = 0.28$, $g_2 = 0.24$, and $g_3 = 0.48$; eq. 8) and the sample-size determination for 1979 where the aggregate quota was 41,610, a table of risk was constructed (i.e., the probability of closing the season before the quota was reached) (Table 3). This table indicates the effect of sample sizes for CV_i = 1.0, 1.5, and 2.0 and CV($\hat{T}_i$) = 0.10, 0.15, 0.20, and 0.25.

For example, when the coefficient of variation of the mortality data within stratum CV_i is 1.5, a sample size of 143 is needed to arrive at $\hat{T}_i$ with CV($\hat{T}_i$) = 0.15 (Table 1). The probability that the estimated total mortality ($\hat{T} = \sum_{i=1}^{3} \hat{T}_i$) exceeds

Table 2. The predicted annual kill of dolphins based on historic (1971–76) yellowfin tuna catch on dolphin and kill-per-ton.

	Yellowfin catch (short ton)	Kill-per-ton	Predicted kill
Max	132,000	0.31	40,920
Min	99,000	0.25	24,750
Mean	117,000	0.28	32,760

the quota depends on the true total mortality. The numerical study indicates that if the true mortality is below 30,000, the estimated mortality rarely exceeds the quota of 41,610 for sample sizes greater than 40. If the true mortality is 40,000, the sample size has to be greater than 140 to guarantee the probability to be less than 25%. Because the dolphin mortality has been monitored by species (stock) since 1977, the P values were also computed for 1 major species of offshore spotted dolphin (*Stenella attenuata*) involved in the fishery. The quota in 1979 was 28,400 spotted dolphins. For a given sample size, the P value for a particular species is greater than that of the aggregate total mortality (Table 3).

DISCUSSION

The sample size determination used in this paper differs from the conventional approach for stratified sampling schemes in 2 respects: (1) coefficient of variation (CV) is used rather than standard deviation (σ), and (2) sample size for each stratum (n_i) is determined and the sum of n_i is the total sample size for the entire year (n).

Basing sample size on the coefficient of variation is preferred for the situation where standard deviation and mean are highly correlated because CV is independent of mean level whereas σ varies with

Table 3. Probability of estimated dolphin mortality ($\hat{T}$) reaching the quota when the true mortality is below quota of aggregate (41,610) and spotted (28,400) dolphins.

$CV(\hat{T})$ $CV(\hat{T}_i)$		0.06 0.10			0.09 0.15			0.12 0.20			0.15 0.25	
CV_i r[a]	1.0 143	1.5 189	2.0 223	1.0 89	1.5 143	2.0 172	1.0 59	1.5 106	2.0 143	1.0 41	1.5 78	2.0 106
True mortality (T)												
20,000 (A)[b]		*			*			*			*	
14,000 (S)		*			*			*			*	
25,000 (A)		*			*			*			*	
17,500 (S)		*			*			*			*	
30,000 (A)		*			*			*			0.01	
21,000 (S)		*			*			*			0.01	
35,000 (A)		0.00[c]			0.02			0.06			0.10	
24,000 (S)		0.00			0.04			0.09			0.15	
40,000 (A)		0.25			0.34			0.37			0.40	
28,000 (S)		0.40			0.44			0.45			0.46	

* $P \leq 0.001$.
[a] See Table 1.
[b] A = aggregate, S = spotted dolphins.
[c] $0.001 < P < 0.005$.

the mean and could be misleading. For example, estimates from 2 years with equal CV are considered to be of the same precision yet their standard errors could be entirely different because of the possible difference in their mean levels.

When the estimated total mortality over time is compared with the quota, the estimates within each stratum are as important as the overall estimate. Normally the sample size (n) is computed according to the precision of the overall estimate $CV(\hat{T})$ for the entire year. The sample size for the ith stratum (n_i) obtained by allocation of n could be inadequate for the precision of within-stratum estimate $CV(\hat{T}_i)$ because $CV(\hat{T}_i)$ is greater than $CV(\hat{T})$ (eq. 8). Therefore, to obtain a desired $CV(\hat{T}_i)$, the sample size for each stratum (n_i) should be determined first as opposed to the conventional procedure by which allocation of sample size to each stratum is performed after the sample size is determined.

This procedure of determining sample size is recommended for estimation problems where the standard deviation is proportional to the mean and the precision of the estimates within each stratum is as important as that of the overall estimate. The allocation method based on coefficient of variation rather than on variance is more realistic and generally applicable to the sampling of natural populations where mean and standard deviation are highly correlated.

Received 25 August 1981.
Accepted 4 July 1982.