Effects of dolphin group type, percent coverage, and fleet size on estimates of annual dolphin mortality derived from 1987 U.S. tuna-vessel observer data

Elizabeth F. Edwards
Christina Perrin
Southwest Fisheries Science Center, National Marine Fisheries Service, NOAA
P.O. Box 271, La Jolla, CA 92038

Abstract.–Amendments enacted in 1988 to the U.S. Marine Mammal Protection Act of 1972, coupled with changes during the 1980's in the composition of the international purse-seine fleet fishing "on dolphin" for yellowfin tuna in the eastern tropical Pacific (ETP), motivated a simulation experiment to assess the effects of fleet size and observer coverage level on estimates of cumulative annual mortality for dolphin stocks, where data quality differed greatly between stocks. Results from this simulation study provided a basis for subsequent U.S. legislation regarding criteria for comparing dolphin mortality rates of U.S. vs. non-U.S. fleets fishing in the ETP, pursuant to regulations governing import of tuna from this area.

Accuracy and precision of mortality estimates were influenced very strongly by data quality (dolphin group type) and moderately strongly by fleet size and level of observer coverage. Because neither fleet size nor dolphin group type can be controlled by a sampling program, desired levels of accuracy and precision in estimates of dolphin mortality can be achieved only by manipulating the level of observer coverage. The unexpectedly strong effect of dolphin group type implies that to protect all groups equally, observer levels should be chosen to accommodate the dolphin group type exhibiting the poorest-quality data.

Manuscript accepted 4 June 1993.
Fishery Bulletin 91:628–640 (1993).

U.S.-registered tuna purse-seiners in the eastern tropical Pacific Ocean (ETP) inadvertently kill dolphins during fishing operations (DeMaster et al., 1992). Since 1972 this kill has been monitored by scientific observers accompanying seiners on routine fishing trips. The kill data are monitored to determine whether annual quotas have been reached for dolphin mortality overall, and for several specific stocks of dolphins. If any quota is reached, fishing must cease on the affected stock for the remainder of the calendar year. Since 1989, the data have also been used to determine annual dolphin mortality rates for comparison with non-U.S. fisheries[1].

Prior to 1987, this kill monitoring was accomplished by observers on 30–50% of the trips made by U.S. purse-seiners during a given year. Thus, mortality was estimated rather than measured directly. In the absence of any alternative (i.e., higher rates of observer coverage), the accuracy and precision of these estimates were assumed to be adequate. This assumption was acceptable at the time because the U.S. fleet was large and generated many data. About 100 vessels fished the eastern tropical Pacific Ocean (ETP) each year, each vessel making two to five trips of about three-months duration each.

The situation changed dramatically in the 1980's when, for a variety of reasons (Sakagawa, 1991), the U.S. fleet began to decrease 20–30% per year, decreasing from over 90 vessels in 1981 to about 10 vessels in 1992. This decrease in size of the U.S. fleet was offset by a reciprocal increase in number of non-U.S. vessels fishing in the ETP. This increase has been dominated by Mexican vessels (fleet size exceeding 50 vessels by 1990) but also includes smaller fleets from several other nations. These smaller fleets range in size from 1 to about 20 vessels (IATTC Annual Reports, 1980–1991). In addition to these changes in composition of the international fleet, amendments passed in 1988 to the U.S. Marine Mammal Protection Act stipulated that yellowfin tuna caught in the ETP could be imported from non-U.S. countries only if mortality rates for those countries were comparable to U.S. kill rates.

These events created a strong incentive to evaluate the effects on mortality estimates of varying observer coverage levels and fleet sizes. Because the new regulations pertained to some individual stocks in addition to dolphin mortality overall, the previously uninvestigated effect of dolphin group type was also of interest.

[1] The mortality estimator used in these simulations (kill/day) is not the same estimator used currently (1993) to determine mortality comparability (kill/set). The reasons for this appear in the methods section.

The U.S. data set for 1987 provided a unique opportunity to investigate these effects. That year, for the first time, observers accompanied nearly 100% of trips by U.S.-registered purse-seiners to the ETP. By subsampling the 1987 data set, various observer levels and fleet sizes could be simulated and their effects on dolphin mortality estimates studied. In the absence of any other available data, these data were taken by NMFS as roughly representative of the international fleet. The simulation results guided subsequent legislation regarding mortality comparability criteria for non-U.S. fleets in the ETP.

Total annual mortality estimates are presented here for three dolphin groups (all stocks combined): spotted dolphin (*Stenella attenuata*); whitebelly spinner dolphin (*Stenella longirostris longirostris*); and common dolphin (*Delphinus delphis*). These three groups represent the range of data types (good, intermediate, and poor in quality) in the 1987 data set, and are generally representative of data types collected every year that may be affected by variable fleet sizes and levels of observer coverage in future fishing years.

Accuracy and precision of the mortality estimates are presented in terms of relative bias and coefficient of variation. Frequency distributions and analytic confidence intervals are examined for underlying patterns in mortality data responsible for generating observed patterns in accuracy and precision.

Methods

Data

During 1987, 34 U.S.-registered purse-seine vessels carrying certificates that permitted fishing "on dolphins" made a total of 134 fishing trips to the eastern tropical Pacific Ocean (ETP). Of these vessels, 33 made 124 trips that were observed by either a National Marine Fisheries (NMFS) or an Inter-American Tropical Tuna Commission (IATTC) technician. These 124 observed trips comprised the entire population of trips for this study. They accounted for 93% of all trips and 96% of all days spent at sea in the ETP by U.S.-registered tuna purse-seiners during 1987. Trips made by individual vessels within fleets were considered replicates and observed trips were assumed to represent unobserved trips accurately. Data consisted of records of date, time, trip number, set number, and kill per set.

Between 4 and 20 vessels carrying observers departed each month during 1987 (Fig. 1). Approximately the same number (although not necessarily the same vessels) of observed vessels (about 30) fished in the ETP during any given month (Fig. 2). Just over half the trips (65/120) lasted about three months. Trip duration ranged from one to six months (Fig. 2). Number of trips incurring mortality of each dolphin group type ranged from 93 (offshore spotted dolphins; Fig. 3) to 12 (common dolphins).

Dolphin mortality data were not stratified geographically, despite the demonstrated efficiency of this stratification scheme (e.g., Hall and Boyer, 1986) because it can be applied only when data sets provide reasonable numbers of reasonably homogeneous data per stratum. Data tend to be so variable for the less frequently killed groups, such as common dolphins, that mortality estimates tend to be very imprecise (and infrequent) even without stratifying data. This is likely to be a problem for all dolphin types for small fleets in every year[2].

Data quality differed between group types in both frequency and extent of kill. Data for offshore spotted dolphins were the "best" in terms of being most plentiful and most evenly distributed between trips collecting data. The frequency distribution of kill per day per trip (i.e., number of dolphins killed per number of days at sea, for a given trip) is relatively smooth (Fig. 3). Most trips killed about 0.5 dolphins per day, a few

[2]At the time this study was conducted (1988–1989), there were no legal requirements for data stratification. Subsequent to this study, the U.S. Congress enacted legislation (Federal Register Notice Vol. 54, #43, Tuesday March 7, 1989; 54 FR9438-9451) requiring stratification of data for annual mortality estimates into three geographic strata and two dolphin groupings (common dolphins, and all other dolphins).

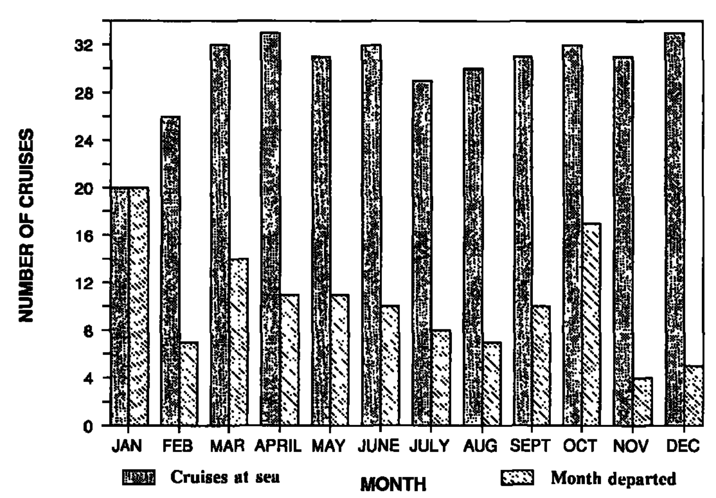

Figure 1
Number of observed U.S.-registered tuna purse-seine vessels at sea, and number of observed vessels leaving port each month during 1987.

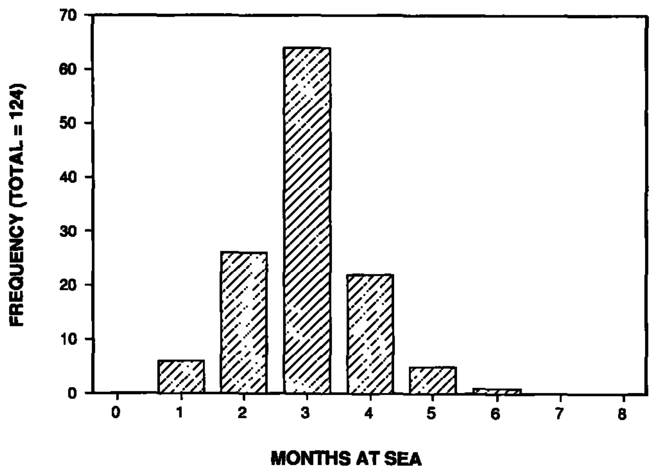

Figure 2
Months spent at sea per trip by U.S.-registered tuna purse-seine vessels carrying observers during 1987.

trips killed no spotted dolphins, and trips with higher kills occurred with decreasing frequency. Data for whitebelly spinner dolphins were intermediate in quality. They were fewer but still relatively smoothly distributed. About half the trips killed no whitebelly spinner dolphins, but of those incurring mortality, most trips were responsible for about 0.1 dolphins per day. Higher kills again occurred with decreasing frequency, with little evidence of extreme outliers. Data for common dolphins exhibited the worst quality, being both sparse and very unevenly distributed. Of the 12 trips incurring mortality of common dolphins, 9 killed about 0.5 dolphins per day, 2 trips killed 1–1.5 dolphins per day, while the remaining trip was responsible for about 9 deaths per day. That one trip was responsible for 67% (594/882) of U.S.-caused common dolphin deaths during 1987.

Simulations

We first selected randomly and without replacement 5, 10, or 20 vessels from the total of 33 observed vessels (Fig. 4). We then sampled randomly and without replacement, 25%, 50%, or 75% of the trips made by each of these selected vessels, without regard to actual timing of those trips throughout the year. We did not force temporal stratification on these small fleet simulations because fleet sizes of 5 or 10 vessels generated so few data.

Vessel selections were replicated 10 times at each combination of fleet size and percent coverage. Trip selections were replicated 50 times from each of the 10 sets of selected vessels. Thus, selected vessels were the same within each set of 50 replicate samples of trips but differed between sets of 50 replicates. Trips conducted by selected vessels in each sample fleet represent the "population" of trips for that fleet replicate.

Each set of 50 replicate selections of trips from a given set of selected vessels generated 50 estimates of cumulative annual mortality (cumulative mortality through December 31). For each set of 50 estimates, we calculated the mean estimate, the coefficient of variation for the mean estimate, and the relative bias of the mean estimate. Thus for each combination of fleet size and percent coverage we generated 10 averaged estimates of annual mortality, 10 coefficients of variation, and 10 estimates of relative bias.

Because the "true" kill varied between fleet replicates depending on the particular vessels and trips selected, the average of these 10 averaged estimates (from the 50 replicates) of annual mortality is not a particularly useful measure in this study. This average represents only the average of the 10 sets of vessels that happened to be chosen for the 10 replicates of fleet size. However, the coefficients of variation and estimates of relative bias for each of the 10 averaged estimates *are* relevant indicators of the ability of the estimation process to precisely or accurately, or both, reflect the true kill, whatever it happens to be for a particular sample of vessels. Accordingly, we present results only for the averages of the 10 estimates of CV and RB generated for each combination of percent coverage, fleet size, and dolphin mortality type. We do not discuss directly the individual mortality estimates.

Estimates

The estimates in this simulation study were derived by using the ratio estimator kill per day, rather than the more precise estimator kill per set (Lo et al., 1982; Hall and Boyer, 1986) because at the time this study was designed and executed, kill per day was still the estimator of choice for NMFS[3].

[3]Kill/day was used by NMFS up until 1987 when 100% observer coverage began, because the quota system required that dolphin kills be monitored continuously throughout the year rather than simply summed at the end of the year. Estimates based on kill/day can be made on a reasonably real-time basis because the day that a vessel leaves port is a readily available datum, and observers report kills by radio on a biweekly basis throughout each observed trip. Number of days fished by the entire fleet (observed plus unobserved) can then be determined relatively simply by summing the number of days at sea since leaving port, and total kill estimated as the product of total fleet days times the estimator kill/day. Biweekly estimates based on kill/set are not possible because observers are not permitted to report set data over the radio. Set data do not become available until after vessels have returned to port.

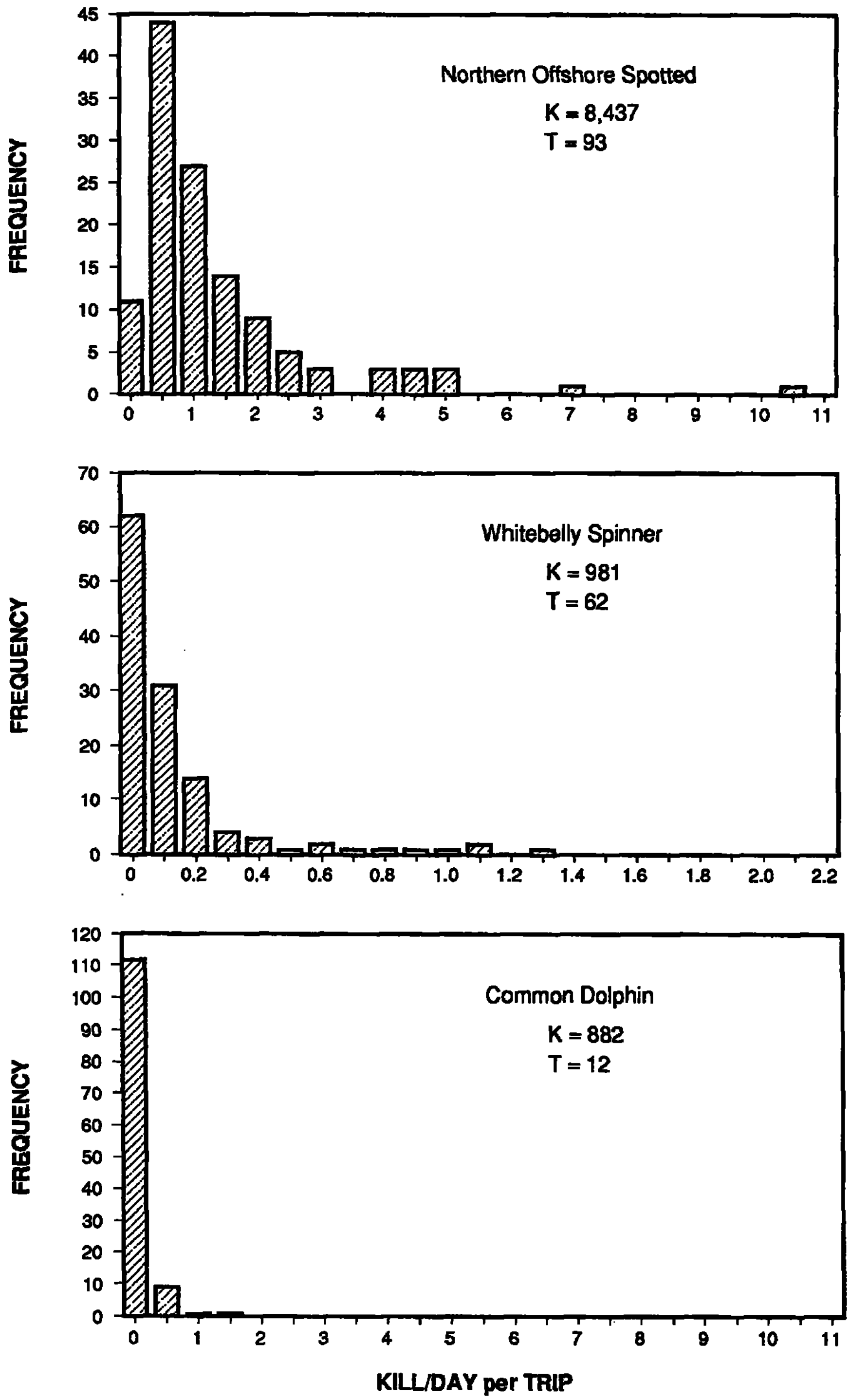

Figure 3

Kill per day during observed trips (124 trips total) by 33 U.S.-registered tuna purse-seiners fishing in the ETP during 1987, for three species of dolphins representing the range of mortality data types. Scales of ordinates and abscissas differ between panels. K is total observed kill during 1987; T is number of trips incurring kill.

Factors and levels included dolphin group$_i$ (i=1,2,3; offshore spotted, whitebelly spinner, and common dolphin), fleet size$_j$ (j = 1,2,3,4; for 5, 10, 20, or 33 boats), fleet replicate$_k$ (k=1,..,10), percent coverage$_l$ (l=1,2,3; for 25%, 50%, and 75%) and percent coverage replicate$_m$ (m=1,...,50).

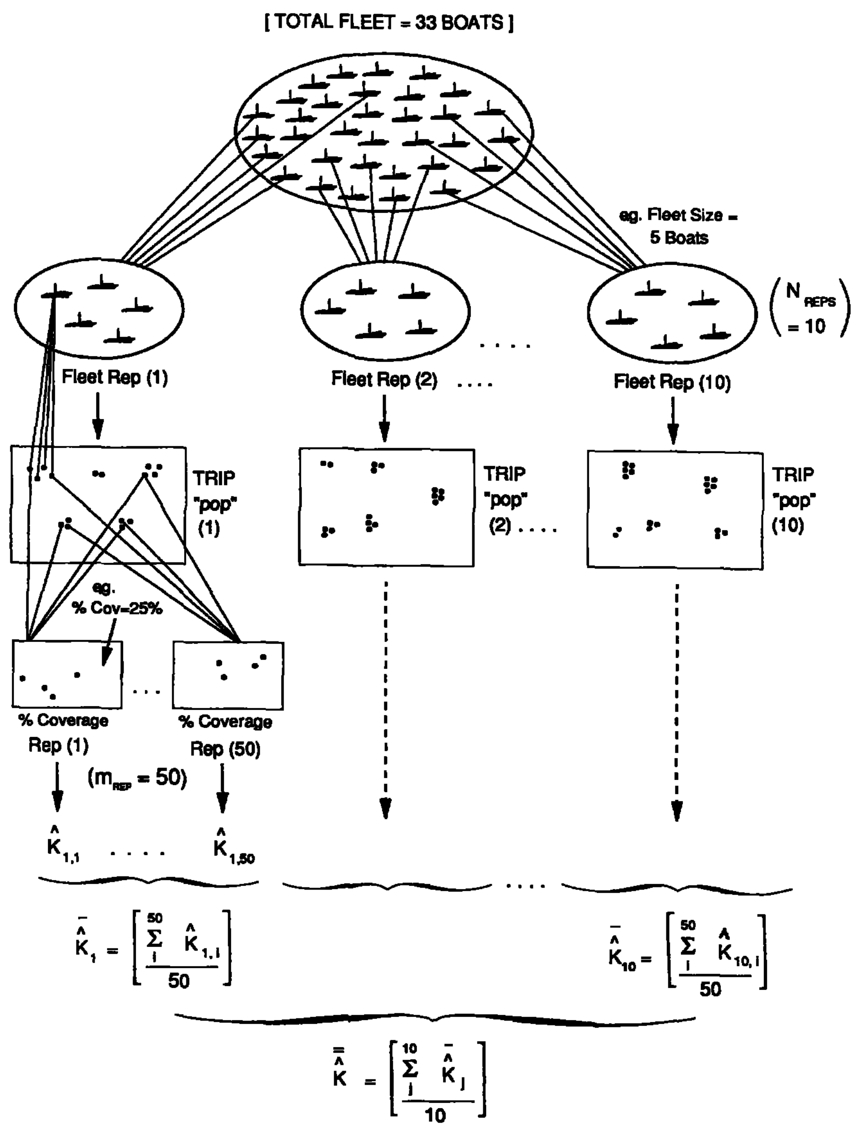

Figure 4
Sampling procedure for simulation study.

Within each replicate, cumulative annual mortality for a given group of dolphins (*YHAT*) is estimated as

$$YHAT_{ijklm} = (K_{ijklm} / DAYS_{jklm}) * TFDAYS_{jk}$$

where K_{ijklm} is the observed cumulative annual kill on all selected trips, $DAYS_{jklm}$ is the observed number of days, and *TFDAYS* is the total number of days spent at sea by all vessels selected for that replicate.

For each set of 50 replicate estimates of annual mortality for a given fleet size, percent coverage, and dolphin group type, we calculated the average estimate ($YAVE_{50}$), the relative bias of that estimate (RBY_{50}), and the coefficient of variation of that estimate (CVY_{50}) as follows

$$YAVE_{50ijkl} = (1/50) * \sum_{m=1}^{50} YHAT_{ijklm}$$

$$RBY_{50ijkl} = 100 * (YAVE_{50ijkl} - known\ kill_{ijk}) / known\ kill_{ijk}$$

and

$$CVY_{50ijkl} = [\sqrt{MSE(YAVE_{50ijkl})} / known\ kill_{ijk}] * 100$$

where

$$MSE(YAVE_{50ijkl}) = \sum_{m=1}^{50} (YHAT_{ijklm} - known\ kill_{ijk})^2$$

and

$$BYAVE_{50ijkl} = YAVE_{50ijkl} - known\ kill_{ijk},$$

where $BYAVE_{50ijkl}$ is the bias of $YAVE_{50}$, and $V(YAVE_{50ijkl})$ is the variance of $YAVE_{50ijkl}$, calculated as

$$V(YAVE_{50ijkl}) = (1/49) * \sum_{m=1}^{50}(YHAT_{ijklm} - YAVE_{ijkl})^2$$

(Cochran, 1977).

The preceding equations produce 10 $YAVE_{50}$'s, 10 RBY_{50}'s, and 10 CVY_{50}'s for each combination of fleet size, percent coverage, and dolphin group type.

The average of these 10 RBY_{50}'s is

$$RBY_{10ijk} = (1/10) * \sum_{k=1}^{10}(RBY_{50ijkl}),$$

and the average of the 10 CVY_{50}'s is

$$CVY_{10ijk} = (1/10) * \sum_{k=1}^{10} CVY_{50ijkl}.$$

Sampling distributions and confidence intervals

To facilitate interpretation of patterns seen in relative bias and coefficients of variations, we plotted both frequency distributions and analytic 95% confidence intervals for an arbitrarily selected single set of 50 individual replicate estimates of mortality derived under various combinations of conditions. The frequency distributions and confidence intervals illustrate, in particular, variations due to differences between replicates in the fleet selected, in contrast to the relative bias and coefficient of variation, which pertain to sampling properties of the estimator. Only one set of 50 replicates, of the 10 sets generated under each combination of fleet size and coverage level, is illustrated for each combination because the general messages conveyed by the figures were the same for all sets.

Frequency distributions and confidence intervals are plotted only for the cases of 5 and 20 boats at 25% and 75% coverage. Combinations of these values spanned the range of fleet sizes investigated and enabled us to examine whether problems might occur even with coverage as high as 75% when fleet size is as small as 20 (or worse, 5) boats.

The frequency distributions of the 50 estimates illustrate graphically the influence of various combinations of fleet size and observer coverage level on the behavior (dispersion) of the estimator itself (kill/day).

The analytic confidence intervals for the 50 estimates illustrate the influence of various combinations of fleet size and observer coverage level, on the estimated precision that may be associated with any individual estimate.

The analytic confidence intervals for individual replicate estimates of mortality were calculated by the International Mathematical and Statistical Library (IMSL) routine SMPRR for ratio estimates (IMSL, 1987). This routine calculates confidence intervals for ratio estimates using the analytic formula for approximate variance of a ratio. The procedure is based on the assumption that a normal approximation to the ratio variance is appropriate (Cochran, 1977; IMSL, 1987). Where data are sparse (fewer than 30 data points in the data set; i.e., in most of the cases in these simulations) this assumption is generally inappropriate (e.g., Cochran, 1977), but for single replicates we had no computationally simple alternative. Bootstrapping confidence intervals for these individual replicates would have eliminated any need for a normal approximation but would have required significantly more computer time to convey essentially the same gross patterns and general message.

Results

Relative bias (RB)

RB was generally small overall, but exceeded the management objective of 5% for common dolphins and whitebelly spinner dolphins when coverage was low (25%; Fig. 5).

Coefficients of variation (CV)

CV decreased with increasing percent coverage and with increasing fleet size in both the "best" data group (offshore spotted dolphin; Fig. 6) and in the "intermediate" data group (whitebelly spinner dolphin). CV decreased with increasing percent coverage but showed no consistent effect of fleet size in the "worst" data group (common dolphin).

Sampling distributions

Frequency distributions were affected somewhat by fleet size (primarily by shifting the central tendency), noticeably by percent coverage (primarily by decreasing the spread of the distribution), and very strongly by dolphin group type (primarily in terms of the number of modes in the distributions; Figs. 7, 8, and 9). Bias and variability increased with small sample sizes and non-smooth data distributions.

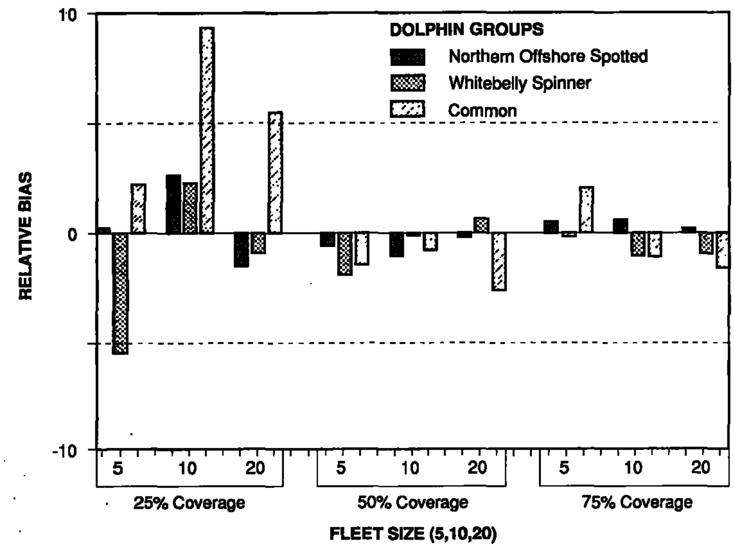

Figure 5
Average relative bias (%) in estimates of annual dolphin mortality as a function of fleet size and percent observer coverage. Dashed line at +5% and −5% indicates management target.

Distributions of mortality estimates for offshore spotted dolphins tended to be relatively narrow and unimodal for fleets of both 5 and 20 vessels, at high coverage (75%; Fig. 7). Distributions remained unimodel but were more dispersed, at low coverage (25%). Distributions for whitebelly spinner dolphins were also unimodal in general but tended to be more dispersed than was the case for offshore spotted dolphins (Fig. 8). Distributions for common dolphins were markedly bimodal and dispersed under all sampling conditions, reflecting the selection (or not) of the one trip with unusually high kill (Fig. 9).

Modal values of mortality estimates for all three dolphin types increased with increasing fleet size (not surprisingly, because more boats generally kill more dolphins) regardless of coverage level. Dispersion also increased with fleet size, more obviously when observer coverage was low than when coverage was high, as more data became available for analysis.

Confidence limits

Confidence limits were affected similarly to frequency distributions. Limits were affected most strongly by dolphin group type, very noticeably by percent coverage, and less strongly but still noticeably by fleet size (Figs. 10, 11, and 12).

Mortality estimates were most stable and confidence intervals dramatically reduced for offshore spotted dolphins at high (75%) observer coverage (Fig. 10). Mortality estimates were more variable and confidence intervals wide and ragged, even with a relatively large fleet (20 boats), for offshore spotted dolphins at low observer coverage (5%). Intervals were most unstable at the combination of lowest observer coverage and smallest fleet size[4]. Confidence interval patterns for whitebelly spinner dolphins were intermediate, being more variable than patterns for northern offshore spotted dophins but being less variable than patterns for common dolphins (Fig. 11). At low coverage and small fleet size, confidence intervals for whitebelly spinner dophins showed the same bimodality characteristic of confidence intervals for common dolphins under all conditions. At high coverage, confidence intervals showed the same relatively stable and

[4]The irregular pattern of the confidence intervals observed for mortality estimates from small fleets (e.g., 5 boats) with low coverage (e.g., 5 boats, 25% coverage) are not surprising because this combination of fleet size and percent coverage means that the mortality estimates are being derived from about 4 observed trips during an entire year (assuming 5 boats make 3 trips/year, 15 trips × 0.25 = 4 trips).

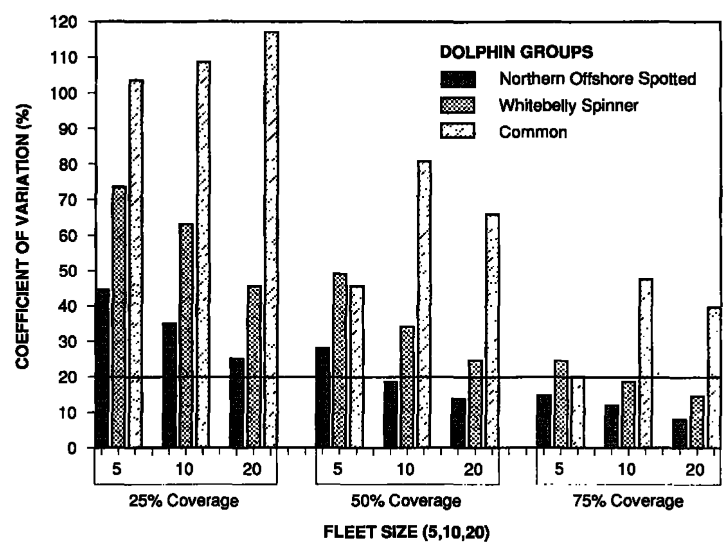

Figure 6
Coefficient of variation (percent) in estimates of annual dolphin mortality as a function of fleet size and percent observer coverage. Dark horizontal bar at 20% indicates management target.

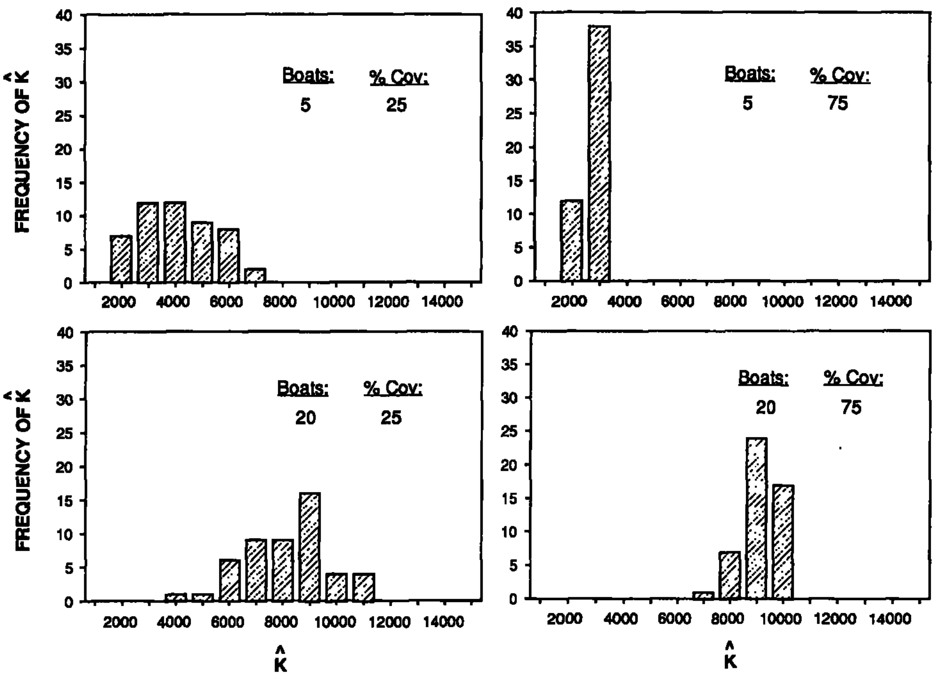

Figure 7

Effect of percent coverage and fleet size on frequency distributions of annual mortality estimates for **offshore spotted dolphins**, from one set of 50 replicates.

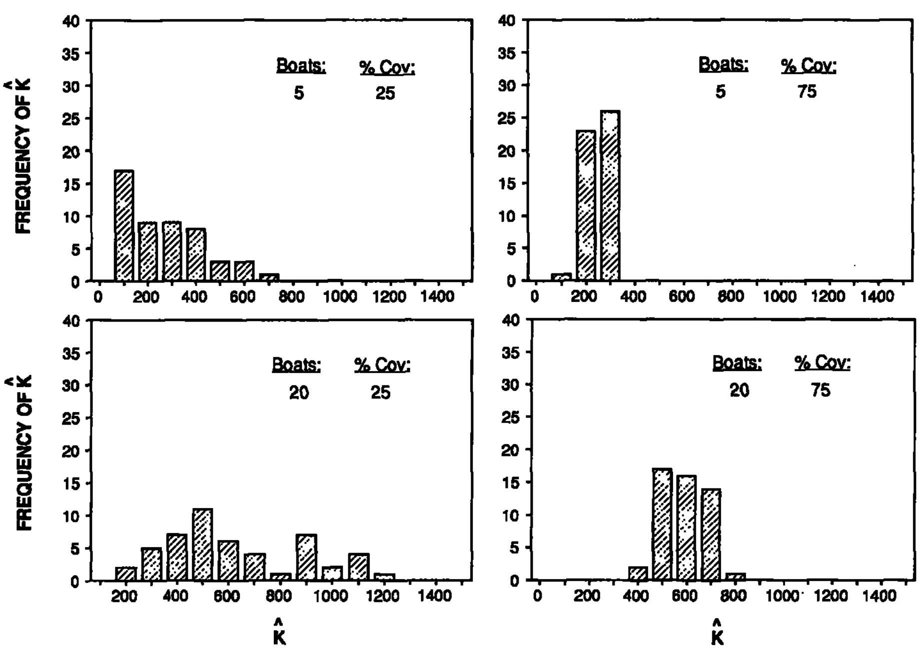

Figure 8

Effect of percent coverage and fleet size on frequency distributions of annual mortality estimates for **whitebelly spinner dolphins**. from one set of 50 replicates.

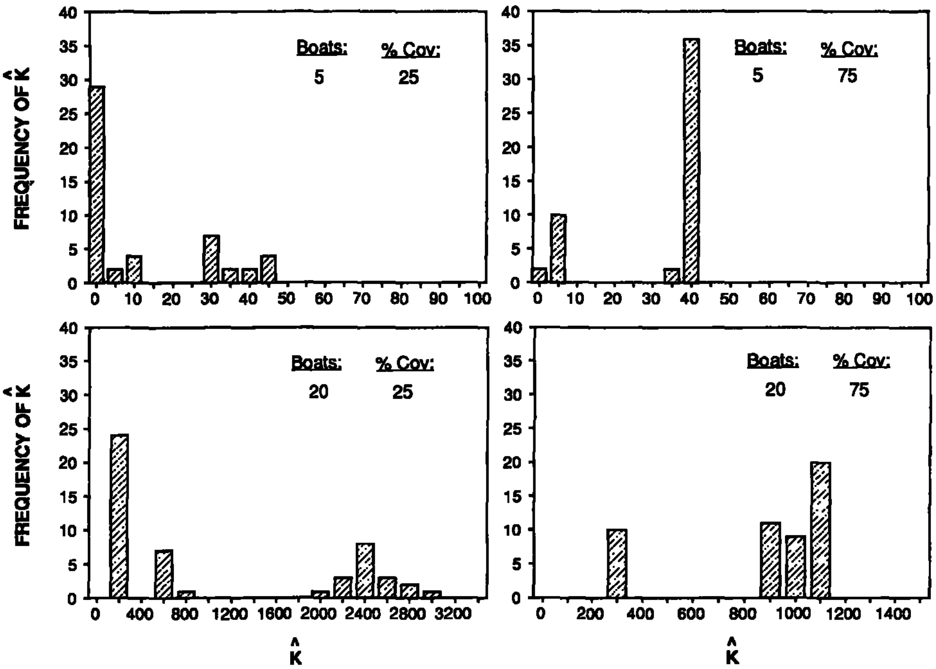

Figure 9

Effect of percent coverage and fleet size on frequency distributions of annual mortality estimates for **common dolphins**, from one set of 50 replicates.

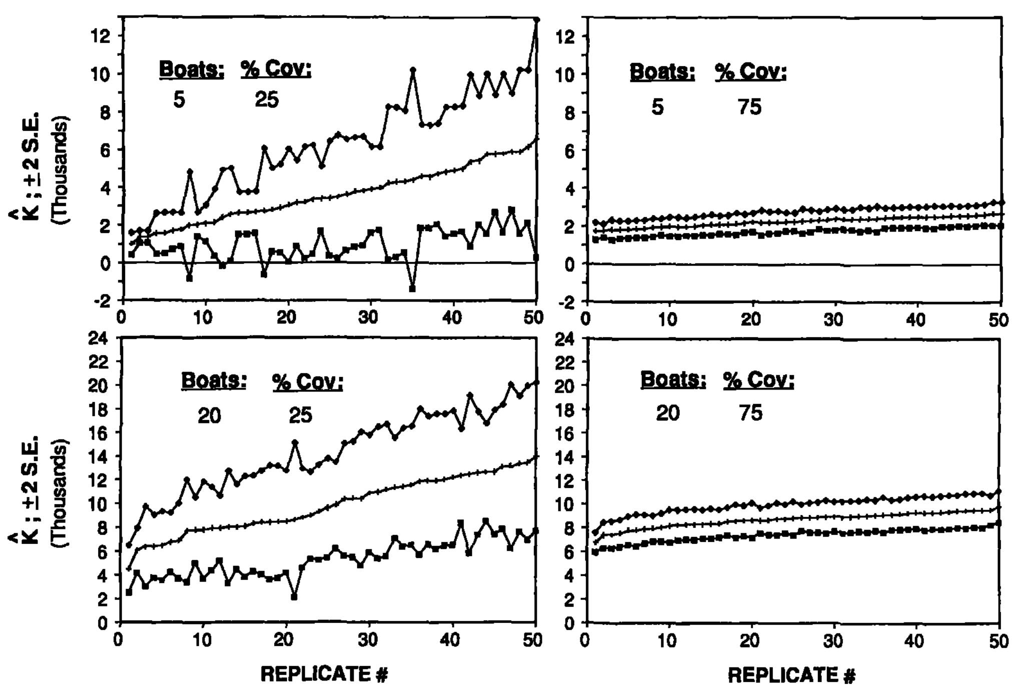

Figure 10

Effect of percent coverage and fleet size on analytic confidence intervals of replicate annual mortality estimates for **offshore spotted dolphins**. Replicate estimates sorted by estimate level, from one set of 50 replicates. Note differences between panels in Y scale.

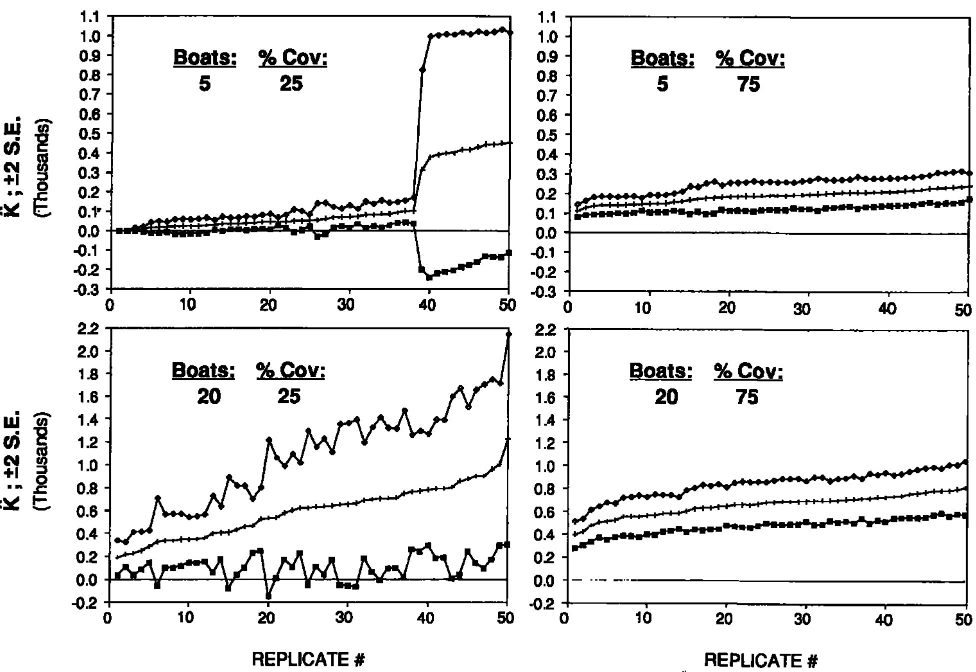

Figure 11

Effect of percent coverage and fleet size on analytic confidence intervals of replicate annual mortality estimates for **whitebelly spinner dolphins**. Replicate estimates sorted by estimate level, from one set of 50 replicates. Mortality estimates expressed in thousands. Note differences between panels in Y scale.

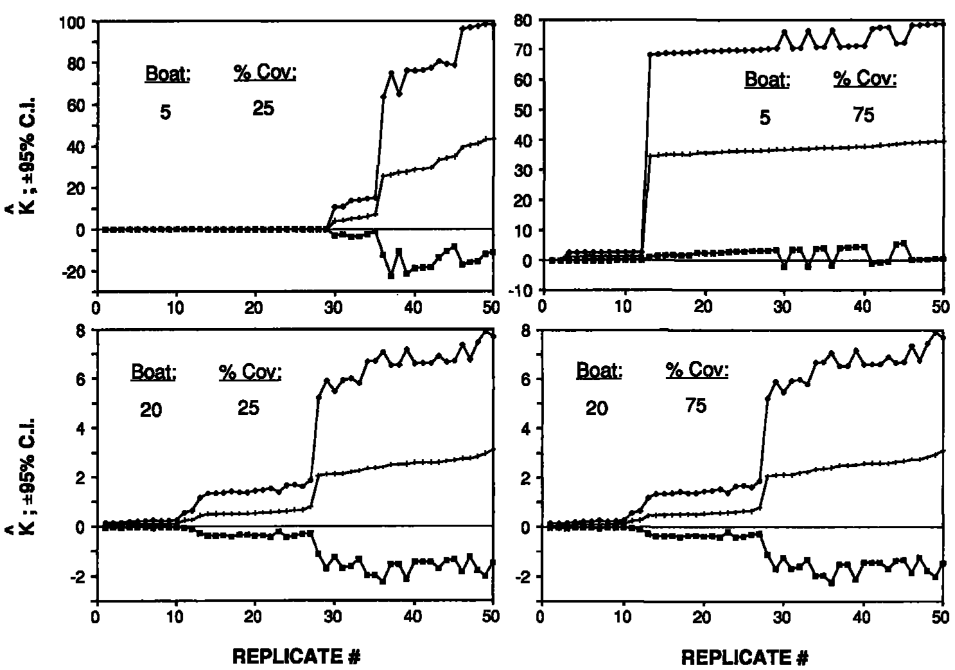

Figure 12

Effect of percent coverage and fleet size on analytic confidence intervals of replicate annual mortality estimates for **common dolphins**. Replicate estimates sorted by estimate level, from one set of 50 replicates. Mortality estimates expressed in thousands. Note differences between panels in Y scale.

narrow limits seen for offshore northern spotted dolphins at high coverage. Confidence intervals for common dolphins illustrate the problems inherent in relatively heterogeneous data, where mortality is sometimes quite high, usually relatively low, and data overall are relatively few. For this species, confidence intervals were extremely variable even with relatively high coverage (Fig. 12) and were very narrow (or nonexistent) for replicates which fortuitously included only low-kill data, but very wide for replicates including a few very high kills.

Although these analytic confidence intervals are clearly inappropriate measures for precise estimation of variance characteristics of the estimates (note negative intervals in some cases; Figs. 10, 11, and 12), they are presented here because they provide an effective illustration of the effects of varying conditions on the variability of the mortality estimates. Despite the inter-group differences, the general response to increasing coverage and increasing fleet size is similar in all three dolphin group types. Confidence intervals become narrower and more stable as more data become available.

Discussion

Dolphin group type

Of the three factors investigated here (dolphin group type, observer coverage level, and fleet size), dolphin group type had the greatest effect on dolphin mortality estimates, followed by percent coverage and fleet size. This hierarchy of effects is controlled by two characteristics of the kill data for each dolphin group type—*frequency* (the number of times that mortality occurs) and *variability* (differences between times in the number of dolphins killed). The *total* number killed can have relatively little influence on the quality of the estimate. This is illustrated by comparing results for whitebelly spinner dolphin and common dolphin. Although total kill was comparable for both dolphin group types (981 deaths of whitebelly spinner dolphin, 882 deaths of common dolphin; Fig. 3) the data sets differed markedly both in number of trips incurring kill (62 for whitebelly spinner dolphin, 12 for common dolphin) and in the distribution of kill per day among those trips (Fig. 3).

The data set for common dolphins exhibits the worst of both characteristics; frequency of kill was low (few trips killed common dolphins) and variability between trips in kill per day was high. These problems exemplify an unfortunate interaction between data collection problems and the ecology of the dolphins themselves. Mortality of common dolphins due to the U.S. fleet during 1987 was *infrequent* because the geographic range of this species is relatively limited and occurs primarily within Mexico's Exclusive Economic Zone. U.S. vessels rarely fish in this area, therefore common dolphins rarely die in U.S. tuna nets in this area. Mortality was *variable* at least in part because common dolphins have an unfortunate habit (in this context) of forming very large schools, pre-disposing them to the possibility of very large-kill "disaster" sets.

The data set for whitebelly spinner dolphin exhibits a problem with only one of the characteristics; data are relatively infrequent. Unlike the case for common dolphins, kill per day was not extremely variable. This similarity in kill per day generates statistics for whitebelly spinner dolphin that are much less biased and variable than for common dolphin.

Observer coverage level

The effect of observer coverage level is influenced both by fleet size and dolphin group type. To achieve a desired level of precision and accuracy in mortality estimates, observer coverage levels will have to be higher in smaller fleets because the available data will be fewer, and higher in dolphin groups with "messy" data, because observer coverage levels affect the probability of encountering an unusually large kill[5].

Observer coverage will need to be relatively high even for large fleets, when estimating mortality of dolphins with sparse and heterogeneous data. For example, with kill data as sparse and variable as was the case for common dolphins in 1987, nearly 100% coverage would be required to generate CV's lower than 20% regardless of fleet size (Fig. 6). With more frequent and less variable kill data, such as for northern spotted dolphin in 1987, CV's lower than 20% can be achieved with 50% coverage of 10-boat fleets (Fig. 6). When fleet size drops to 5 boats, even this relatively well-behaved data set requires coverage at about 75% to achieve CV's less than about 20%. With coverage as low as 25%, even a fleet size of 20 boats was insufficient to meet the management objective of 20% CV.

Fleet size

In smaller fleets, each data point comprises a larger fraction of the available mortality data. In particular, the influence of unusually large mortalities (e.g., the

[5]The possibility of high kills is more important than the possibility of low kills; this is because all dolphin group types experience zero kill frequently (therefore it is not unusual or unexpected). Also, zero kill is a definitive lower bound, while the upper bound on kill is limited only by the potential school size of the dolphin group.

one trip responsible for over 500 deaths of common dolphins) will be much greater in smaller fleets.

The accuracy and precision of mortality estimates are affected less by fleet size *per se* than by observer coverage level for a given fleet size, and the fleet's variability in kill per day. Kill per day is affected not only by the fishermen's choices of fishing methods and areas, but also by the type of dolphin found associated with a given school of tuna. Because these are not factors that can be controlled, small fleets will generally require higher coverage level than larger fleets to achieve a given level of accuracy and precision in mortality estimates.

Simulation procedures and estimates

The simulation procedure used in this study was designed to reflect the sampling process as it would occur in the real world. More precise and less biased estimates of mortality rates for the population of trips contained in the 1987 data set would have resulted from simple random sampling of the 124 trips in the data set as a whole. But simple random sampling implicitly assumes that all vessels are equal in fishing ability. This is not the case, and it is likely that some fleets as a whole may have greater (e.g., those newer to purse-seining and therefore less experienced) or lesser (e.g,, the more experienced fleets) mortality rates than the average for the ETP purse-seine fleet overall. In addition, in the real world, not all trips made by all purse-seine vessels fishing in the ETP would be available for sampling. Only trips made by the vessels in a particular fleet would be available for sampling, and only those vessels actually observed would contribute data. If vessels (or more properly, the crew) differ in their ability to release dolphins unharmed (or not), then fleets with more (or fewer) "low kill" vessels will have lower (or higher) mortality rates than other fleets of comparable size. Although it would have been possible to estimate the number of trips that would have been made, on average, by a fleet of a given size, and to have then randomly sampled that many trips from the 1987 data base, the results would have been unrealistically precise. The cluster sampling resulting from the selection of trips only after selecting vessels adds variability in the estimates but is more realistic than simple random sampling. The sampling scheme used here is a single stage cluster sampling, for which a ratio estimator is the most appropriate choice of estimation procedure (Cochran, 1977).

Discussion

Although the kill-per-day estimator used in this simulation study is no longer used by NMFS because 100% observer coverage has made estimation unnecessary, the results of this study have general implications for current estimation procedures based on kill per set (e.g., Hall and Boyer, 1986) and for mortality estimation procedures in general where data quality may vary between stocks.

The uneven structure of the data set for common dolphins has unfortunate implications for deriving estimates of mortality for dolphin groups that are characterized by having such infrequent and widely variable kill per day. Specifically, estimates of mortality can vary widely depending on which trips happen to be chosen. In our simulation, we could resample the total population of vessels repeatedly, thus generating relatively unbiased, though individually variable, estimates of mortality. In the real world, only one sample (one set of mortality data per dolphin group type) will be collected per year. If this sample is collected under low percent coverage, it appears very likely that the data may be affected by undetectable sampling biases. This bias is more likely to underestimate than to overestimate mortality because sets with large kill are rare and likely to be underestimated, even though the mortality during such sets may be responsible for a disproportionately large percentage of the total kill.

The problem with missing the rare large-kill sets is that the kill in these sets can apparently be one or two orders of magnitude greater than the "usual" kill. For very abundant groups, missing a few large kills will miss only a small percentage of the total number of dolphins in the group; underestimating mortality could be relatively harmless. For less abundant groups, the large kills might represent a significant proportion of the existing stock. Underestimating this mortality could lead to seriously underestimating the impact of mortality due to fishing operations on these stocks.

In the case of the 1987 data set for dolphin kill by the U.S. fleet, coverage greater than 96% (the highest observed) would be required for all boats in order to generate mortality estimates for common dolphin with CV's less than 20%. Alternatively, if only the most abundant groups are considered (e.g., offshore spotted dolphin), CV's less than 20% could be achieved with only 50% coverage of fleets as small as 5 boats. There appears to be no unique solution that is optimal for all groups.

In addition, the poor quality of the data presented here for common dolphin in fact underestimates the true extent of the problem for this species. In actual practice, the species is managed as three separate stocks rather than as one combined stock as presented here. The data are thus extremely sparse for the individual stocks, and the problems with estimating mortality, given anything less than full observer coverage, are greatly exacerbated.

The results of these simulations, in particular the results for five-boat fleets, and common dolphins, strongly influenced subsequent regulations related to comparability criteria for import of tuna caught by non-U.S. fleets. These regulations now require that mortality data for common dolphins be stratified separately from all other stocks, and that non-U.S. fleets meet the U.S. requirement for 100% observer coverage[6].

Of the three factors discussed here (dolphin group type, percent coverage, and fleet size), only percent coverage can be controlled by the sampling program. Our results imply that providing maximum protection for all dolphin groups would require mandating coverage to achieve a desired level of statistical precision for the dolphin group type with the least statistically stable data.

Acknowledgments

The U.S. fleet simulation study was designed by Bruce Wahlen. Nancy Lo and Peter Perkins provided insightful advice about sampling procedures and statistical methods. We thank also the participants in the Workshop on Estimating Dolphin Mortality, and two anonymous reviewers, who provided many useful comments during the review process.

[6]U.S. regulations for acceptable levels of observer coverage for non-U.S. fleets have progressed from a requirement for 33% coverage of trips made by fleets with 10 or more vessels, 50% of trips on fleets of 5–9 vessels, and 100% of fleets with 1–4 vessels during 1989 and 1990 (Federal Register 12/19/89;Vol. 54 #242, p. 51918–51923), to 75% of all trips regardless of fleet size during 1991 and 1992 (10/18/90; 55 FR 42235 and 10/8/91; 56 FR 50672), to the current requirement for 100% coverage regardless of fleet size (Federal Register 1/8/92; Vol. 57, #8, p. 668).

Literature cited

Cochran, W. G.
1977. Sampling techniques, 3rd ed. John Wiley & Sons, NY, 428 p.

DeMaster, D. P., E. F. Edwards, P. Wade, and J. E. Sisson.
1992. Status of dolphin stocks in the eastern tropical Pacific Ocean. *In* D. R. McCullough and R. H Barrett (eds.), Wildlife 2001: populations. Elseiver Press, London, England.

Hall, M. and S. Boyer.
1986. Incidental mortality of dolphins in the eastern tropical Pacific tuna fishery: description of a new method and estimation of 1984 mortality. Rep. Int. Whal. Comm. 36:375–381.

IATTC Annual Reports.
1980–1991. InterAmerican Tropical Tuna Commission, c/o Scripps Institute of Oceanography, La Jolla, CA 92037, var. pagination.

IMSL.
1987. SMPRR/DSMPRR. *In* IMSL, Inc. Stat. Library, International Mathematical and Statistical Library of Computer Programs, p. 751–758.

Lo, N. C. H., J. E. Powers, and B. E. Wahlen.
1982. Estimating and monitoring incidental dolphin mortality in the eastern tropical Pacific tuna purse seine fishery. Fish. Bull. 80(2):396–400.

Sakagawa, G. T.
1991. Are U.S. regulations on tuna-dolphin fishing driving U.S. seiners to foreign-flag registry? North Am. J. Fish. Mngmnt. 11(3):241–252.