HOW MUCH OBSERVER COVERAGE IS ENOUGH TO ADEQUATELY ESTIMATE BYCATCH?

Elizabeth A. Babcock
Ellen K. Pikitch

Pew Institute for Ocean Science
Rosenstiel School of Marine and Atmospheric Science
University of Miami
4600 Rickenbacker Causeway
Miami, FL 33149-1098

Charlotte G. Hudson

Oceana
2501 M Street, NW, Suite 300
Washington, DC 20037

Summary

Bycatch is a critical source of mortality for marine species, including endangered species, heavily fished commercial and recreational target species, and many species of so-called "trash fish" whose importance in marine food webs is now being recognized. Whether management objectives include conservation or fisheries yields, adequate measurement of at-sea mortality is a necessary component of any management framework, and observers at sea are the most reliable source of information. The amount of observer effort, when not financially constrained, is usually set to achieve a desirable level of precision, assuming that the observers sample the fleet randomly. The assumption of random sampling is often unjustified, as the sampling process is both conceptually and operationally very complex. The issue of bias in bycatch estimates is often not addressed, despite the fact that many observer programs allocate sampling effort opportunistically to vessels that volunteer to carry observers. The bias introduced by non-random sampling, and by the changes in fishermen's behavior in the presence of observers, must be addressed. Comparing the catches of observed and unobserved vessel-trips should be an ongoing component of any observer program. If the observer samples are an unbiased sample of the fishery, our literature review and simulation studies suggest that coverage levels of at least 20 percent for common species, and 50 percent for rare species, would give reasonably good estimates of total bycatch. The required level of coverage, however, could be much higher or much lower for a particular fishery, depending on the size of the fishery, distribution of catch and bycatch, and spatial stratification of the fishery.

Table of Contents

Introduction

This paper addresses the question of which level of observer coverage is sufficient to estimate total bycatch with enough accuracy and precision[1] for management. Observer coverage refers to the fraction of fishing effort (e.g. vessel trips) that is sampled at sea by trained scientific data collectors. Several definitions of bycatch and discards can be found in the fisheries scientific literature. We will use the definitions of Hall: "Capture" is everything that is caught during fishing operations; "catch" is the part of the capture that is retained; "bycatch" is the fraction of the capture that is discarded dead—whether it is a protected animal like a dolphin, a non-marketable fish, or a fish of the target species that is undersized; and "release" is the portion of the capture—protected species or fish—released live (Hall 1996, 2000).[2]

When designing an observer sampling program the level of coverage required will depend on the objectives of the observer program, which might vary from estimating bycatch of protected species, to improving bycatch and catch data for assessment of fish populations, to collecting biological data. This report focuses on the *statistical issues in estimating bycatch* of protected species, incidentally caught species, or individuals of target species that must be discarded for any reason. Although the legal requirements for estimating bycatch of marine mammals are different from the requirements for estimating bycatch of commercially valuable species, the statistical issues are similar.

This report also focuses on bycatch of *individual species*. In designing observer programs, it is common for several species to be of interest—undersized fish of a target species, a non-target species that is important in another fishery, or a protected sea bird or mammal, for example. In such cases, the sampling requirements for each relevant species will depend on the species frequency of occurrence, patchiness, seasonality, variability in recruitment, and other factors. Fisheries scientists should design sampling programs that are adequate for all species of interest.

Is 100% observer coverage necessary?

In some cases, particularly where low levels of mortality may jeopardize the recovery of a threatened or endangered species, it may be necessary to have an exact count of the total incidental mortality. In such cases, 100% observer coverage is necessary. Endangered species interactions have prompted requirements for 100% observer coverage in several U.S. fisheries, such as the Atlantic shark gillnet fishery, during times of the year when right whales are calving (NMFS 2002b).

---

[1] There are two sources of error in discard estimates taken from observer sampling programs: accuracy, which measures how close the expected value of the estimate is to the actual value, and precision, which measures how close a series of independent estimates are to each other.
[2] Other definitions exist. Alverson (1994) defined bycatch as dead discards, plus live releases, plus retained catch of species that were not targeted (also known as incidental catch). It is also common to refer to "discards" of fish and shellfish, and "bycatch" of mammals, turtles and birds.

Endangered species interactions are not the only argument for 100% coverage. It is common for coastal nations to require 100% observer coverage on foreign vessels fishing within the domestic exclusive economic zone (EEZ). This is partly an enforcement measure and partly a recognition that more accurate total catch data improves management (see references in Nolan 1999). Some purely domestic fisheries also require 100% coverage because they are managed with sophisticated in-season management measures as well as state-of-the-art assessment methods that require accurate and timely data about catch and bycatch. The Alaska groundfish fishery, which requires levels of coverage of 100% for vessels greater than 124 feet in length and of 30% for vessels 60 to 124 feet in length, is a good example. These levels of observer coverage were set more than ten years ago to ensure sufficiently precise estimates of catch and bycatch composition for assessment and management (NMFS 2002a, Table 1). Over time, the high quality of the data available in this fishery has given scientists and managers more options for developing in-season management measures such as bycatch quotas (see Ackley and Heifetz 2001). In particular, for a fishery to be managed with individual vessel quotas on bycatch, 100% observer coverage is generally necessary. This is the case in the eastern tropical Pacific tuna purse seine fisheries, in which the Inter-American Tropical Tuna Commission requires 100% coverage so that individual vessel quotas on dolphin bycatch may be used (IATTC 2003).

If observer coverage is less than 100%, what level is sufficient?

If a level of 100% observer coverage is not attainable, then the coverage level chosen must ensure that the total bycatch estimate is sufficiently accurate and precise for assessment and management purposes. The precision of an estimate depends on the size of the sample, the size of the fishery, and the variability of the bycatch. The accuracy of an estimate depends on these three measurements as well as whether the sampled part of the fishery is representative of the entire fishery. Determining the sample size necessary to estimate a quantity (such as total bycatch) with a desired level of precision is one of the fundamental problems inherent in the scientific method, and thus the appropriate calculations that need to be made are well known (see Cochran 1977, Rao 2000). Some characteristics of fisheries data, such as the temporal and spatial variability of fisheries and the difficulty of random sampling, decrease the likelihood of achieving precise estimates. The rate at which unwanted fish or other animals are caught—and the rate at which they are discarded—is influenced by many factors. Some of those factors may change over time, leading to a necessary change in the level of coverage over time as well. For example, changes in the relative abundance of target and non-target species, changes in market prices, short-term changes in winds or currents, and changes in regulations or the enforcement of regulations— along with characteristics of the vessels related to gear and fishing operations— all complicate sample-size calculations. More importantly, estimates of total bycatch from observer data can be biased (i.e., not accurate) if the coverage is less than 100%.

A considerable amount has been written about how to develop appropriate stratification schemes and other statistical methodologies to achieve the best estimate of total bycatch for a specified level of sampling effort (see Cotter 2002, Allen et al. 2002). In addition, various papers on estimating appropriate sample sizes for observer programs, looking only at the issue of precision under the assumption that sampling is random and representative, can be found in the peer-reviewed literature and in fisheries agency "gray literature" (see Allen et al. 2002, Dinardo 1993, Hay et al. 1999, Fogarty and Gabriel 2002, Karp and McElderry 1999). Unfortunately, there have been few papers examining the critical issue of potential bias in observer data.[3] Those papers that address the issue, however, show that bias can be significant, particularly at low-coverage levels. Simulation studies by Hall (1999) showed that bias was quite high in total catch estimates of a dolphin species in the eastern tropical Pacific purse seine fishery at sampling levels of 10% or less, even when using bootstrap methods to correct for bias.

This report will discuss some of the possible sources of bias in bycatch estimates from observer sampling programs as well as how these problems can be solved. We will also discuss some issues related to precision. We will not go into great detail about the many statistical estimators for total bycatch and the variance of total bycatch, or about statistical methods such as bootstrap bias correction, since these estimators and methods are well documented in the fisheries and statistical literature.

What causes bias in observer estimates of bycatch and how can such bias be reduced?

While the issue of precision is considered in the design of most observer sampling programs, most of these programs either fail to consider bias or dismiss it as insoluble (see Fogarty and Gabriel 2002). Nevertheless, it is possible to test for bias using available data, and methods can be developed to minimize bias. Simply increasing sampling fractions can reduce many types of bias in observer data.

What causes estimates of total bycatch from observer data to be biased? Observer data are very reliable for the portion of the fishery that is observed. But if the observed areas are not representative of the fishery as a whole, bias may be introduced when the bycatch estimated from observer samples are extrapolated over the rest of the fishery to estimate total bycatch. Observer samples will not be representative of the fishery if, for example, 1) bycatch rates change when observers are on board, 2) voluntary vessel participants have different bycatch rates than non-participants, or 3) logistical constraints are related to bycatch rates (Fogarty and Gabriel 2002, Liggens 1997). Bias can also be caused by inaccurate recording of data by observers, by small sample sizes, and by inappropriate stratification.

[3] For example, in the *Aquatic Sciences and Fisheries Abstracts* (ASFA), the most comprehensive abstracting service for the fisheries literature, there were 98 references between 1978 and the end of 2002 that reported on observer bycatch data. Of these, only seven mentioned the subject of bias or accuracy in their abstracts (Buchary 1996, Byrne and Pengilly 1990, Cotter et al. 2002, Edwards and Perrin 1993, Medley 2001, Wahlen and Smith 1985, Walsh et al. 2002). These papers, and the few others that addressed bias and did not come up in the ASFA search (Liggens et al. 1997, Sampson 2002, Hall 1999), cover only a few of the many observer programs in existence.

Bias caused by "observer effects"

The presence of an observer on board a vessel can cause the vessel crew to change the decisions they make about where to fish, which species to target, how to configure the fishing gear, and which species to discard (Hall 1999, Liggens 1997). Consequently, bycatch rates estimated from observed trips may not accurately reflect bycatch rates of the fleet as a whole. For levels of coverage below 100%, this "observer effect" can bias the estimates of total bycatch from observer studies.

It is possible to measure whether observer-effect bias exists. In southeastern Australia, for example, Liggins et al. (1997) compared estimates of total landed catch from observer data to the weight of the landed catch for a multispecies fish trawl fishery. The observer estimates were expanded to the entire fishery using the known total fishing effort. In some cases, measured total catches were outside the confidence bounds of the observer estimate, but the size and direction of the error varied. In other words, there was no consistent bias in observer estimates of landings. Liggens et al. (1997) also found no bias in the size distributions of the observer estimates of landed catch versus the port sampled catch. They concluded, therefore, that there probably was not an observer-effect bias. This result greatly increased the credibility of the total bycatch estimates from this observer program and demonstrates that methods to determine whether observer data are biased do exist.

In other fisheries, observer-effect bias may be more pronounced. In the Oregon groundfish trawl fishery, for example, Sampson (2002) examined the species composition of landings from observed and unobserved trips. He found that trips that carried an observer were significantly different than trips that did not, implying that the observer data were not representative of the fleet as a whole.[4] This bias in the observer data could have been the result of different fishing behavior in the presence of observers, or of a non-random allocation of observers across fishing trips. As Sampson (2002) pointed out, these results imply that total estimates of bycatch based on observer data may not be reliable. In a very different fishery—the Alaska crab pot fishery—Byrne and Pengilly (1997) compared estimates of target species catch from observer data to landings data, including information from confidential interviews with fishermen, to determine whether the observer data were accurate. For components of the fishery with high coverage the observer estimates were quite good, but in the Adak brown king crab fishery, which had 12% coverage, the relative error of the observer catch-per-unit-effort (CPUE) estimate was an unacceptable 35%. The authors attributed this error to the fact that the sampling effort was not well distributed across the large area and long fishing season of the brown king crab fishery.

Estimates of total bycatch from observer studies have also been compared to estimates of total bycatch from fishermen's logbooks for several fisheries (see Brown 2001, Walsh et al. 2002) and differences have been found. It is generally not clear, however, whether the differences are

---

[4] Sampson (2002) performed a multivariate analysis called PCA (principal components analysis) on the weight of landings of various species by vessel trips. The PCA results from trips that carried an observer were significantly different than the PCA results from unobserved trips, implying that the observer data were not representative of the fleet as a whole.

caused by changes in fishing behavior in the presence of an observer or logbook misreporting, or both.

These studies show that bias can exist, and that it is possible to measure the level of bias, at least in landed catch. Biases in the species composition, length-frequency or total weight of the landed catch, or in the spatial and seasonal distribution of sets, provide some indication of whether the bycatch estimates are likely to be biased. Clearly, it is necessary to monitor for observer-effect bias, using methodology like that of Liggens et al (1997) and Sampson (2002). Also, higher levels of observer coverage would reduce this bias by increasing the fraction of trips that are observed (and thus leading to bycatch rates accurately estimated by the observer program). And if there were a financial cost to the fishermen for fishing differently when observers are on board, then higher levels of coverage would provide an incentive to fishermen to fish in a more typical manner.

Bias caused by non-random allocation of sampling effort

In any scientific sampling program, samples must be taken randomly so that they will be independent of each other even while being from the same statistical distribution. Such independent, identically distributed (i.i.d.) samples are necessary for the sample to be expanded correctly to estimate totals, means, and variances for the population (Cochran 1977, Rao 2000). In observer data, random sampling is complicated by the fact that observers take a sample of the fishery at multiple levels; they sample a fraction of vessels, on which they sample a fraction of trips, during which they sample a fraction of fishing operations, from which they sample a fraction of the catch or bycatch (Tamsett and Janacek 1999). Sampling must be random at every level, with appropriate coverage to achieve precise estimates and avoid sampling bias. A random sample of fishing operations is very difficult to achieve, and even a random sample of vessels can be problematic. For example, Liggens et al. (1997) reported that some skippers refused to carry observers and some were simply difficult to contact when the observer program was trying to arrange trips, which led to a non-random allocation of sampling effort. If sampling is allocated randomly to trips and trips vary in length, then tows might not be randomly sampled because the longer trips will have more tows observed. Finally, if the number of boats fishing is difficult to quantify, then randomly allocated sampling effort may be problematic because, without knowing the total number of boats in each component of the fishery, it would be difficult to assess whether the samples are representative (Liggens et al. 1997).

In observer programs with less than 100% coverage, it is common for observers to be placed on vessels that volunteer and have the space and facilities to accommodate an observer. Samples taken in this opportunistic fashion are quite likely to be biased by the fact that vessels that are willing and able to carry observers on board may not be representative of the fishery. To avoid bias caused by non-representative sampling, observers must be allocated randomly or systematically across the fishery.[5] Voluntary observer programs should be avoided, but if such programs are inevitable, then comparing the spatial distribution and landed catch weight, length

[5] Systematic sampling (for example, every 10th trip) can be easier to implement than random sampling. Systematic sampling can complicate the calculation of variances of discard estimates, but there are methods available to help avoid this problem (Conquest 1996).

distribution, and species composition from vessels with observers to those without—as in Liggens et al. (1997) and Sampson (2002)—would give some indication of whether the observed trips are representative of the fleet as a whole. Finally, if vessels that volunteer for the observer program at low levels of coverage have lower discard rates than vessels that are not observed, then increasing the level of coverage might include more vessels that are more typical of the fishery (Hall 1999).

Bias caused by logistical constraints

In some cases, there are components of the fishery that are logistically difficult to sample, leading to biased estimates of total bycatch. For example, smaller vessels may not be safe for observers, or some ports, seasons, or fishery components may present more difficulty in sampling than others (Liggens et al. 1997, Cotter 2002). Observer programs that sample only during part of a fishing year may be problematic if the amount of discarding depends on management measures, such as trip limits, that are adjusted throughout the season (NMFS 2001, Sampson 2002).

As with the observer-effect bias and volunteer bias, the solution to this problem is to make every effort to take a random (or at least representative) sample of the fishery. It is also possible to develop a stratified sampling regime that takes some logistical constraints into account so that the total bycatch estimate will be unbiased. For example, Cotter (2002) used a stratified sampling method called "probability of sampling proportional to size" (PPS) to address this issue. This method involved allocating sampling effort to fishing vessels with a formula that made the probability of sampling a vessel proportional to the size of the vessel and its historic fishing effort, and inversely proportional to the average length of a vessel's fishing trips. The stratified sampling method thus allowed the observers to spend less time on small, inconvenient boats while still allowing an unbiased estimate of total bycatch. Nevertheless, if there are categories of vessels on which it is impossible to place an observer, then the bycatch rates for that type of vessel will not be well estimated, even with increased levels of sampling. In such cases, the uncertainty in estimates of bycatch should be acknowledged when the data are used for science and management.

Bias caused by inaccurate recording of data by observers

Some at-sea observers may deliberately under-report bycatch, due to friendships with the vessel crew, intimidation, or even bribery, which would lead to an underestimate of total bycatch. To determine whether an observer is consistently under-reporting bycatch, the Inter-American Tropical Tuna Commission (IATTC) observer program uses, "a set of statistical procedures, comparing observers among themselves in a set of well-defined strata, to identify those that tend to fall consistently on one side of the distribution" (M. Hall, IATTC, pers. comm.). Such data-checking methods should be common practice in observer programs.

Bias caused by small sample size

Even if the observed trips are representative of the fishery, estimates of total bycatch can be biased when low sample sizes are used. If the statistical distribution of the bycatch is particularly "clumped," meaning that most sets have zero bycatch while a few have very high bycatch, then a small sample size will lead to biased estimates of total bycatch. Bycatch commonly has this sort of distribution. In 2001, for example, 25% of the dolphin bycatch in the eastern tropical Pacific tuna purse seine fishery occurred in a single set (IATTC 2001, M. Hall, IATTC, pers. comm.). With such data, a much larger sample size is needed to get an accurate bycatch estimate.

Small sample bias is common if total bycatch is estimated with a ratio estimator. With a ratio estimator, the average ratio of bycatch to landed catch is estimated from the observer sample, and this value is multiplied by the total landed catch to estimate total bycatch. While ratio estimators generally give more precise estimates of total bycatch than can be achieved with a simple sample (Saila 1983), the ratio estimator can be biased at low sample sizes (Cochran 1977, Rao 2000). Various methods to adjust for bias in ratio estimators, including bootstrap bias correction methods (Chernick 1999), have been proposed and are sometimes used (Hall 1999 and references therein). The level of bias caused by small sample sizes can be estimated for a particular fishery by using simulation studies. For example, Hall (1999) reported on a study of the biases of discard ratio estimates of dolphin bycatch in tropical tuna fisheries.[6] That study showed that—for the dolphin-tuna fishery data—all of the ratio estimation methods demonstrated high levels of bias at low sampling fractions (below 20%), although bootstrap bias correction methods greatly improved the estimates (Hall 1999).

The problem of bias caused by low sample sizes is commonly ignored in observer program sampling design, but can be solved by increasing the sampling fraction. Also, simulation studies similar to Hall (1999) should be used to test the proposed estimators of total bycatch, and to develop estimation methods that are unbiased for the fishery being sampled.

Bias caused by inappropriate stratification

Observer samples are usually stratified by quarter, gear type, fishing area, and other factors, thus increasing the precision of total bycatch estimates for a given level of coverage. The stratification scheme allows more sampling effort to be allocated to sectors of the fishery that have more variable bycatch, so that the estimates of total bycatch will be more accurate and precise (see Allen et al. 2002, Cochran 1977, Conquest 1996). By modeling bycatch in the French trawl fleet in the Celtic Sea, for example, Rochet et al. (2002) concluded that greater observer sampling effort in sectors of the fishery with more variable discarding behavior would improve the precision of the total bycatch estimates. An appropriate stratification scheme can

[6] The observer data from three consecutive years was taken to be the "universe" and various levels of sampling coverage were randomly sampled from this known universe. Several methods for estimating discard ratios were used, and the relative bias in the ratios calculated by each method was plotted for various levels of sampling coverage (Hall 1999).

greatly increase the precision of total bycatch estimates for a given level of sampling effort. Bias can be introduced into the total bycatch estimates, however, if some strata are not adequately sampled.

For stratified sampling designs to provide precise and unbiased estimates of total bycatch, each stratum should have a sample size of at least twenty to thirty observations (Hall 1999, Cochran 1977, Conquest 1996). A sample size of at least three is needed to estimate a variance for the stratum (Dinardo 1992), and if some strata are assumed to have zero variance—when in fact the variance is not zero—the total variance will be underestimated (Fogarty and Gabriel 2002, Pikitch and Babcock 2002). The amount of coverage necessary to adequately sample small strata can be surprisingly high. For example, the U.S. Atlantic pelagic longline fishery has a target coverage level of 5% systematically allocated across the fleet (Brown 2001), but for the majority of the strata sample sizes were so low that the bycatch rates had to be estimated by pooling strata.[7] For the New England groundfish fishery, the overall coverage level was 1-2%, but for many species and fishery components, there was not enough data to estimate a variance (Fogarty and Gabriel 2002). In assessments, the strata are often collapsed[8] to deal with these sample size issues, or discard-to-kept ratios from adjacent strata are applied to strata for which there are no data (Brown 2001, Fogarty and Gabriel 2002). These ad hoc methods complicate variance estimation and can introduce positive or negative bias if not used with caution.

Even with high levels of observer coverage randomly distributed across the fishery, the under-sampling of small strata (i.e., minor components of the fishery) can lead to biased estimates of total bycatch. To demonstrate this, we simulated fisheries data with three strata, one of which included only 10% of the total fishing trips (see Appendix 1 for details). We created an observer sampling program with coverage levels that ranged from 1% to 50%, and repeated the observer sampling process 500 times at each level. The results are illustrated in Figure 1, which shows how the percent error from the true value of total bycatch varies with the amount of observer coverage for a small fishery (100 trips per year) and a larger fishery (1000 trips per year). For the large fishery, even the smallest stratum had an adequate sample size at low coverage levels. The small fishery demonstrated the potential negative bias caused by under-sampling small strata. At low coverage levels, the estimated bycatch rate in the smallest stratum was often zero because the few sampled tows happened to not have bycatch, leading to low bycatch estimates at low coverage (Fig. 1). Thus stratification schemes should be designed so that no stratum is too small to be adequately sampled.

Bias caused by small samples in minor strata can be easily avoided by stratifying fisheries in such a way that each stratum will have a sample size of at least 20-30. At low levels of coverage in a small fishery, stratification may not be possible. Higher levels of observer coverage would allow finer stratification schemes to be used.

---

[7] For 1999, there were 420 sampled sets, but only 5 of the 24 quarter/area combinations in the stratified sampling design had 30 or more observations. Thus, for 19 of the 24 strata, the bycatch rates had to be estimated by pooling strata, first across years, then across quarters.

[8] "Collapsing" strata means recombining areas, seasons, and other factors that had been separated in the original sampling plan to create a new stratum with a larger sample size.

Reducing bias (and increasing precision) by modeling the discarding process

We have discussed how to reduce bias in observer estimates using straightforward statistical sampling, but it is also possible to improve total bycatch estimates by increasing our understanding of discarding practices. Bycatch of potentially marketable fish can be quite variable, depending on the discarding decisions made by individual fishermen. Fishermen may retain a species sometimes and discard it other times, depending on their fishing location, how long they have been away from port, the size and quality of the fish, which species they are targeting, the current prices the processing plants are paying for target and incidentally caught fish, and the regulations in effect at the time. Estimates of total bycatch can be improved—both in precision and accuracy—by understanding and modeling the discarding decisions made by fishermen.

Research has recently been conducted on modeling discard rates (Rochet et al. 2002, Walsh et al. 2002, Gillis et al. 1995, Sampson 2002).[9] In the U.S. Pacific coast groundfish fishery, for example, individual vessel landings are constrained by species-specific "trip limits" that are adjusted throughout the fishing seasons. Changes in trip limits have been shown to change discarding decisions (Pikitch 1991, Pikitch et al. 1998) because they change the economic incentives to discard[10] (Gillis et al. 1995). These types of regulations must be taken into account when estimating total bycatch in order to avoid bias (NMFS 2001, Sampson 2002). Such studies demonstrate that greater understanding of fishermen's discarding decisions can improve the design of observer programs, and can also improve estimates of total bycatch from the fishery, using both observer data and landings or logbook data.

How precise should bycatch estimates be?

Assuming that observers can sample a fishery randomly, and that there is no sampling bias or observer-effect bias, how precise do estimates of total bycatch need to be? The answer depends on how the estimates will be used. In some cases, the level of precision needed in an estimate of total bycatch is quite high because the bycatch species is endangered (e.g., sea turtles off Hawaii, NMFS 2000) or is being targeted by another fishery (e.g., Pacific halibut caught as bycatch in demersal trawl fisheries off Alaska—Karp and McElderry 1999). The required level of precision

[9] Walsh et al. (2002) developed a generalized additive model (GAM) to predict the bycatch of blue sharks in unobserved sets by vessels in the Hawaii-based pelagic longline fishery, by combining observer coverage with oceanographic data and vessel and gear characteristics. Walsh et al. (2002) compared the model's predictions to logbook data to examine logbook reporting practices. Other studies that have used models of discarding behavior to improve observer sampling design or discard estimates include Stratoudakis et al. (1999) in the North Sea gadoid fisheries, Tamsett and Janacek (1999) in the English North Sea fishery, and Rochet et al. (2002) in the French trawl fleet in the Celtic Sea.

[10] The model of Gillis et al. (1995) showed that even with constant total allowable catch limits for the entire Oregon, Washington, and California trawl fishery, the proportion of sablefish capture that would be discarded could vary from less than to 10% to more than 80% with different trip limits.

is less obvious when trying to estimate bycatch for fisheries stock assessments[11] and management because the amount of information needed for management depends on both the assessment methodology to be used and the management system itself. For example, a management system based on marine reserves and gear regulations might not require estimates of total fish biomass, while a system based on harvest control rules would require such information (Walters 1998). Thus an observer sampling program should be designed within the context of a fishery management system in order to support management objectives such as rebuilding overfished populations of target species (Williams and Corral 1999, Rohan 1999). In a fishery where bycatch mortality is high compared to other sources of mortality, higher levels of coverage are needed.

The level of precision needed for a particular assessment and management system can be determined with a simulation study. Punt (1999), for example, developed a simulation to examine the costs and potential benefits to management of various levels of observer coverage in the blue grenadier trawl fishery off the coast of Australia. The approach used a simulated data system to examine the performance of various management strategies related to achieving management objectives under various levels of observer coverage. For management schemes based on estimating the biomass of spawning fish, he concluded that precise observer data were not necessary. He also noted, however, that for management schemes requiring precise estimates of the strength of recent-year classes of juvenile fish, precise observer data greatly improved management.

While this result is specific to the Australian blue grenadier fishery, several general conclusions can be drawn from the study. First, the necessary level of precision for an input to a stock assessment depends on the specific fish, fishery, population dynamics model, and management system being considered. Second, it is possible to achieve an improved understanding of the necessary level of observer coverage through simulation of the system. Finally, the same method could be applied to determine the optimal allocation of research funding to sampling of different sectors—such as commercial or recreational—of a fishery.

Required precision of bycatch data for commercially important species

Accurate measurements of how many fish, including both catch and bycatch, are killed by fishing operations, are a high-priority input for fisheries stock assessment models. Fisheries agencies often gather data on every landed fish by requiring the weights of fish sold to processing plants to be reported, although a sub-sample of the landings would theoretically give a fairly precise estimate of total landed catch. The estimates of bycatch from the observer program are added to these precisely estimated landings data (and to recreational fishery data) to estimate the total fishing mortality for assessment. Although it would be impossible to achieve the same level of precision in bycatch estimates as in catch estimates, the precision of bycatch estimates should be quite high, particularly if the bycatch is large compared to the catch.

[11] "Stock assessment" is the scientific basis of fisheries management. It involves the use of mathematical models, combined with fisheries data, to predict the effect of management measures on fish populations.

As Rice (1999) pointed out, inaccuracies in catch data are often the largest source of uncertainty in stock assessments. One common stock assessment model, virtual population analysis (VPA), estimates the total numbers of fish of each age in a population each year by extrapolating from the catch-at-age (including for bycatch) in each year. In a “tuned” VPA, indices of abundance, such as survey data or catch-per-unit-effort (CPUE) series, are used to adjust the estimates of model parameters so that the total population trend is consistent with the abundance indices. The abundance index data are assumed to have errors, thus the model does not have to fit them exactly. Even in a tuned VPA, however, it is generally assumed that the catch-at-age data are error-free. This means that, in practice, errors in catch data are much more problematic than errors in survey data (Hilborn and Walters 1992, Mohn 1999).

Even sophisticated statistical catch-at-age models, which allow catch to be estimated with error, make the assumption that errors in catch data are not consistent over time. Many stock assessments are subject to “retrospective bias,” meaning that subsequent assessments of the same population vary in a consistent manner (Mohn 1999). It is common for the more recent assessments of a stock to show that the previous assessments were overly optimistic about the stock’s status and ability to rebuild. Such retrospective biases cause management actions to be based on an incorrect understanding of status and resilience of the stock, and are a major cause of failures in fisheries management (Rice 1999, Smith 1998, Mohn 1999). Retrospective bias can be caused by changes in discarding practices over time (Rice 1999). Discarding practices commonly change over time as the prices of various species change and also as regulations change (Babcock and Pikitch 2000, Pikitch et al. 1988, Pikitch 1991). Observer studies are a way to document discarding and ensure that changes in discarding practices over time will not bias subsequent assessments.

In summary, if observer data are used to estimate bycatch of commercial species and the bycatch is high relative to the catch, then the bycatch should be estimated at least as precisely as the catch, unless it can be shown that the precision of the bycatch estimates can be lower without negatively affecting management (Punt 1999).

What sampling fraction is needed to achieve precise estimates of total bycatch?

Assuming that observers can sample a fishery randomly, and that there is no sampling bias or observer-effect bias and that the appropriate level of precision has been determined based on how the bycatch estimates will be used, what is the level of observer coverage needed to achieve a desired level of precision? If the sample really is random, then standard power and sample size calculations apply (see Cochran 1977, Rao 2000). These calculations have been made for many observer sampling programs (see Dinardo 1993, Hay et al. 1999, Fogarty and Gabriel 2002, Karp and McElderry 1999). The necessary sample size can be lower if an appropriate stratification scheme is used to optimize the allocation of sampling effort. There is considerable literature on the development of appropriate stratification schemes to improve precision for a specified level of sampling effort (see Cotter 2002, Allen et al. 2002), and on the selection of appropriate estimators—the statistical formulas used to calculate total bycatch and its variance—(see Allen et al. 2002, Hall 1999, Ortiz et al. 2000).

The appropriate stratification, allocation of sampling effort, and level of sampling to achieve a given level of precision vary depending on the specific characteristics of the fishery. The diversity of fishing practices in the fishery, the rarity and level of aggregation of the bycatch, the number of trips per year, and the kinds of management measures in place all influence the level of coverage required to achieve a desired level of precision. The sampling level needed to adequately sample a fishery can change with time if the abundance or spatial distribution of the target or bycatch species changes, or if the fleet changes its spatial distribution. Also, observer sampling is a multi-level sampling process, and the number of trips that are required to be sampled to achieve a given level of precision will depend on the fraction of sets that are sampled by observers during each trip (Karp and McElderry 1999, Allen et al. 2002). An example of the variability in required coverage levels can be seen in the achievement of an excellent level of relative precision (1-20%) for the various components of the French Celtic Sea trawl fishery, with coverage levels of 0.8-1.5 % (see Rochet et al.). Karp and McElderry (1999), however, cited simulation studies of the Alaska groundfish trawl fishery that showed that, if observer-effect bias did not exist and vessels could be sampled randomly, a coverage level of 30% was required to estimate total bycatch of common species with a "reasonable" level of precision, assuming observers sampled 60-70% of hauls in each observed trip. Rare species required an even greater sampling fraction.

Precision for rare versus common species

Generally, bycatch of a species that is commonly encountered or has low variance can be measured with a lower level of coverage than bycatch of a rare species or a species with highly variable catch rates (Cochran 1977, Hall 1999, Karp and McElderry 1999, Rochet et al. 1998).

To illustrate this, we simulated fisheries in which the bycatch species varied from rare to nearly as common as the targeted species, while the target species catches, sets per trip, and so on were the same in all fisheries (see Appendix 1 for the specifics). We randomly drew 500 observer samples from the fishery under various levels of observer coverage, and determined the level of observer coverage required so that more than 90% of the simulated observer samples would estimate a total bycatch within 10% of the correct value (Table 3a, Figure 3). Because the estimated bycatch is being compared to the actual known total bycatch, this measure of error includes both accuracy and precision. The simulated data do not include any observer-effect bias, but there may be minimal sample bias (see bias section above). The total bycatch was estimated with a combined ratio estimator (Appendix 1). For a rare species (bycatch less 0.1% of catch), the required level of coverage to achieve the specified level of accuracy and precision was more than 50%, while around 17% coverage was sufficient for the common species (bycatch 35% of catch, Figure 3, Table 3a). While the exact level of coverage required for a particular fishery would depend on the distribution of the discard and catch species, it is clear that, in order to accurately assess bycatch, rare species require substantially higher levels of coverage than common species.

Because the required level of sampling differs for different fisheries, observer programs should simulate an observer sampling program from their own preliminary data to determine the required level of coverage for accurate and precise estimates. Pilot studies should be done to estimate the expected catch and variability of the bycatch or discard species, so that this data can be incorporated into the sampling design.

Precision in large versus small fisheries

We commonly talk about levels of observer coverage in observer program design, but the precision of a statistical estimate is generally calculated as a function of sample size, not sampling fraction. The terms "sample size" and "sampling fraction" are often used interchangeably since, if the size of a fishery is known (e.g., in number of trips per year), it is easy to convert sampling fraction to sample size:

sample size = sampling fraction x number of trips

It is sometimes assumed that the precision of an estimate depends only on the sample size, and that sampling fraction does not matter (Fogarty and Gabriel 2002). In other words, a sampling fraction of 10% of trips in a fishery with 1,000 trips per year would be exactly equivalent to a sampling fraction of 1% of trips in a fishery with 10,000 trips per year. In fact, because observer programs are sampling without replacement from a finite population, a higher sampling fraction will lead to a more accurate and precise result than a lower sampling fraction for the same sample size. In other words, the number of trips in the fishery influences the precision of an estimate.

For example, with the simulated data discussed above, the sampling fraction required to ensure that at least 90% of the simulated observer data sets estimated a total discard within 10% of the correct value was 3.6% for a fishery with 10,000 fishing trips and 28% for a fishery with 1,000 trips. This corresponded to a sample size of 360 trips for the larger fishery and 280 trips for the smaller fishery (Table 3b, Figure 4, Appendix 1). Put simply, although a higher sampling fraction was needed for the small fishery, a higher number of trips sampled was needed for the large fishery. This demonstrates that the size of the fishery matters when calculating the required sample size and sampling fraction for an observer program.

How can simulated data analyses help inform our observer programs?

The following case studies are theoretical simulations of two domestic U.S. fisheries with observer programs. Although the appropriate coverage level for any observer program should be estimated using real data using the methods described in the previous sections (see Hall 1999, Liggens et al. 1997, Sampson 2002), simulations provide a quick "back-of-the-envelope" estimate of the appropriate levels of coverage. We chose to simulate a trawl fishery based on data from the U.S. Pacific groundfish trawl fishery, and a gillnet fishery based on data from Atlantic coastal gillnet fisheries. The species captured as bycatch in these two fisheries ranged from infrequent bycatch of a protected species to bycatch of a finfish that was almost as common as the target species of the fishery (Appendix 1). For each fishery, we randomly drew 500 observer samples from the simulated fishery for various levels of observer coverage, and then determined the level of observer coverage required so that more than 90% of the simulated observer samples would estimate a total discard within 10% of the correct value. The total bycatch was estimated with a combined ratio estimator that took the catches in each trip into account (see details in Appendix 1).

The simulations in each case assumed that the observer samples were representative of the entire fishery, although there is evidence of observer-effect bias in the Pacific fishery (Sampson 2002). Thus the estimates of required sample size from this analysis should be considered minimal estimates based on precision and sampling bias only—as considerations of observer-effect bias would increase the required sample sizes.

Pacific coast groundfish trawl fishery

The U.S. Pacific coast groundfish fishery stretches from Washington to California and includes both shallow and deep-water species. The management plan covers 82 species, which live on or near the sea floor. We simulated trawl data for both the deep-water complex and the shallow-water fishery as similarly as possible to the actual fisheries in the years 2001 and 2002, using preliminary data from the observer program (NWFSC 2003). We simulated the capture of two common species in the DTS (Dover sole, thornyhead, and sablefish) complex North of 40°10'N, as well as a rare species in the flatfish complex all along the Coast. Dover sole and sablefish are commonly found in the deep-water DTS complex and are captured even when untargeted. The bocaccio is a severely depleted Pacific groundfish, managed under a rebuilding plan, that is captured more rarely in the shallow-water trawl fishery but that, by regulation, must be discarded.

For sablefish, the average observed bycatch in each depth and season stratum in the DTS complex ranged from two to 180 pounds per hour (NWFSC 2003). Bycatch rates varied greatly from one tow to the next; the standard errors in the bycatch-per-hour estimates in Table 4 of NMFSC (2003) imply that the CV's[12] of the bycatch per tow were between 130% and 590% by stratum. For this commonly caught species, the sampling fraction required that more than 90% of the simulations be within 10% of the true value was between 0.3 and 0.4 (Fig. 2, Table 2). For Dover sole, the observed bycatch in each stratum was between 9 and 60 pounds per hour (with CV's between 120% and 400%), and the sampling fraction required was between 0.3 and 0.4 (Fig. 2, Table 2).

For bocaccio captured in the flatfish fishery, the observed bycatch rates in each depth, area, and season stratum ranged from 0.009 to 70 pounds per hour (with CV's between 170 and 1130). The number of simulated observer samples that were within 10% of the correct estimate of total bycatch increased more gradually with the sampling fraction because bocaccio bycatch is rare and highly variable (Fig. 2, Table 2). The sampling fraction required for 90% of the simulated observer samples to estimate a total bycatch within 10% of the true value was much higher than 50%.

---

[12] The coefficient of variation (CV) of the catch per hour is a measure of how variable the catches per hour are, equal to the square root of the variance of the catch per hour by tow divided by the mean catch per hour. Ninety-five percent of the sets would have a catch within 2 CV's of the mean catch.

Similar to the theoretical simulations described earlier, these simulated case studies indicate that sampling fractions need to be quite high to accurately account for rare bycatch species. The result that coverage levels of 30-40% are required is consistent with studies done in the Alaska groundfish fishery (Karp and McElderry 1999), and appears to be a result of the high variability in the catches. The required sample fractions in these case studies are higher than those in the theoretical simulations presented earlier, most likely because the assumed variability in the bycatch was higher in these two fisheries, and because more variable catches require higher coverage.

Southeastern Atlantic coastal gillnet fisheries

The Atlantic coastal bottlenose dolphin is a rare but protected species among the bycatch of southeastern U.S. gillnet fisheries. Listed under the Marine Mammal Protection Act as depleted, this species is the subject of a take-reduction team and has specific fisheries regulations aimed at minimizing its occurrence as bycatch. Because the dolphin has been observed caught in all gillnet mesh sizes, diverse fisheries along the Atlantic coastline are affected.

We simulated the bycatch of Atlantic coastal bottlenose dolphins in southeastern coast gillnet fisheries—an extremely rare occurrence (Palka and Rossman 2001). Between 1996 and 2000, only 12 coastal bottlenose dolphin captures were observed in 1,876 observed trips. Because of the extreme rarity of dolphins in the catch, Palka and Rossman (2001) estimated bycatch rates using a generalized linear model (GLM), a method well adapted to extremely rare events (Ortiz et al. 2000). We simulated catch of bottlenose dolphins as a rare event, consistent with the Palka and Rossman (2001) analysis, but estimated total bycatch with the same ratio estimator used in our other simulations for consistency (see Appendix 1). Because the bycatch of bottlenose dolphins is a rare event, the sampling fraction required to achieve bycatch estimates within 10% of the correct value in 90% of the cases was greater than 50% (Fig. 2, Table 2).

Summary and discussion

Observer programs are widely recognized as the best way to obtain reliable information about bycatch and discarding activities that take place at sea. The amount of observer sampling effort is often constrained by the amount of money and other resources available to the program. Nevertheless, for observer programs to provide adequate information to improve fisheries stock assessments, endangered species protections, and ecosystem management, programs should be designed to achieve the objectives of the observer sampling program, which will generally require high or moderate levels of precision and minimal bias in estimates of total bycatch.

The issue of bias in bycatch estimates has not been given the attention it deserves in the design of sampling programs. In particular, many existing observer programs allocate observer sampling effort opportunistically to vessels that volunteer or are willing to carry observers. The bias introduced by non-random sampling must be addressed if the data collected by an opportunistic sampling program is to be at all useful. The work of Liggens et al. (1997) and Sampson (2002) demonstrate that it is possible to determine whether or not an observer program is gathering data that is representative of the fleet as a whole. Such analyses should always be done. The problem of accuracy should not be ignored because it is more difficult to measure bias than precision.

Comparing the catches of observed and unobserved vessel trips (Liggens et al. 1997, Sampson 2002) should be an ongoing component of any observer program, whether or not the observers are allocated randomly. Of course, a mandatory observer program with randomly allocated observers would produce more reliable results.

Once the issues of observer effects and sampling bias have been addressed, the desired level of precision in bycatch estimates should be determined by examining how the observer data will be used in assessment and management. A management strategy evaluation such as that described by Punt (1999) would be ideal but time consuming. It should be possible, however, for the designer of any observer sampling program to find out approximately what level of precision is needed in the bycatch estimates for each species.

Once the appropriate level of precision has been determined, the question of required level of sampling effort can be addressed. Determining the appropriate level of sampling effort is an iterative process. Early in the development of an observer program, when no data are available, the level of coverage could be set by comparison with other observer programs, or by general sample size considerations. Our literature review and simulation studies suggest that coverage levels of at least 20% for common species and 50% for rare species in a fishery with more than a few thousand trips per year would give reasonably good estimates of total bycatch.

If some information is available about the expected rarity, distribution, and variability of the bycatch species in the catch, then it is possible to get a better estimate of the required sample size through a simulation study like that presented above for Pacific groundfish and Atlantic coastal bottlenose dolphins. If some observer data have already been collected for the fishery of interest, then it is possible to simulate an observer program sampling from the actual observed trips (Hall 1999), and to get a very good estimate of the sample size required to obtain precise and accurate estimates of bycatch. Once some observer data have been collected, appropriate use of stratification schemes (see Rochet et al. 2002, Cotter et al. 2002) can reduce the sampling effort (and hence the expense) required to achieve a given level of precision.

Recommendations

Observer programs should:

1. Simulate observer samples from actual data to find coverage levels that estimate bycatch with an appropriate level of precision for assessment and management. Unless managers can show that the lower levels of coverage give sufficient precision and accuracy, we suggest—based on our simulated data applications—that if the bycatch species is rare, observer programs should adopt coverage levels of at least 50%. And if the bycatch species is common, observer programs should adopt coverage levels of at least 20%.

2. Compare landings and spatial and seasonal distribution of observed and non-observed trips to determine whether there is evidence of observer-effect bias. If bias exists, the sampling design must either improve randomization or increase sample size, or both.

3. Determine the level of precision required for discard estimates by examining the sensitivity of the stock assessment models actually used for key species to the precision of the discard estimates.

4. Sample the fishery randomly or systematically and cover all components of the fishery, allocating observer coverage levels high enough to adequately sample every stratum of the stratified sampling design.

Acknowledgements

Thanks to M. Hall and an anonymous reviewer for comments on an earlier draft. Thanks also to M. Hirshfield for review and comment and for helpful discussions in the initial conception of this project. This work was supported by a grant from Oceana.

References

Ackley, D. R. and Heifetz, J. 2001. Fishing practices under maximum retainable bycatch rates in Alaska's groundfish fisheries. *Alaska Fishery Research Bulletin* 8: 22-44.

Allen, A., D. Kilpatrick, M. Armstrong, R. Briggs, G. Course, N. Pérez. 2002. Multistage cluster sampling design and optimal sample sizes for estimation of fish discards from commercial trawlers. *Fisheries Research* 55: 11–24.

Babcock, E. A. and E. K. Pikitch. 2000. A dynamic programming model of fishing strategy choice in a multispecies trawl fishery with trip limits. *Canadian Journal of Fisheries and Aquatic Sciences* 57:357-370.

Brown, C. A. 2001. Revised estimates of bluefin tuna dead discards by the U.S. Atlantic pelagic longline fleet, 1992-1999. *ICCAT Collective Volume of Scientific Paper*s 52: 1007-1021.

Buchary, E. A. 1996. Fisheries catches. *Fisheries Centre Research Report* 4(1) 57-59. University of British Columbia.

Byrne, L. C. and Pengilly, D. 1998. Evaluation of CPUE estimates for the 1995 crab fisheries of the Bering Sea and Aleutian Islands based on observer data. p. 61-74 in Funk, F. and 7 co-editors, *Fishery Stock Assessment Models.* Lowell Wakefield Fisheries Symposium Series No. 15.

Chernick, M. R.1999. *Bootstrap methods: a practitioner's guide.* John Wiley and Sons. New York.

Cochran, W. G. 1977. *Sampling Techniques,* 3rd Edition. John Wiley and Sons. New York.

Conquest, L., R. Burr, R. Donnelly, J. Chavarria and V. Gallucci. 1996. Sampling methods for stock assessment for small-scale fisheries in developing countries. p 179-225 in Gallucci, V. F, S. B. Saila, D. J. Gustafson and B. J. Rothschild, eds., *Stock assessment: quantitative methods and applications for small-scale fisheries.* CRC Press. New York.

Cotter, A. J. R, G. Course, S. T. Buckland and C. Garrod. 2002. A PPS sample survey of English fishing vessels to estimate discarding and retention of North Sea code, haddock, and whiting. *Fisheries Research* 55: 25-35.

Dinardo, G. T. 1993. Statistical guidelines for a pilot observer program to estimate turtle takes in the Hawaii longline fishery. NOAA Technical Memorandum NOAA-TM-NMFS-SWFSC-190.

Edwards, E. F. and Perrin, C. 1993. 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. *Fishery Bulletin* 91: 628-640.

Fogarty, M. J. and W. L. Gabriel. 2002. Relative precision of discard rate estimates for the Northeast groundfish complex. Report of the National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole.

Gillis, D., R. M. Peterman, and E. K. Pikitch. 1995. Implication of trip regulations for high-grading: a model of the behavior of fishermen. *Canadian Journal of Fisheries and Aquatic Science* 52: 402-415.

Hall, M.A. 1999. Estimating the ecological impacts of fisheries: what data are needed to estimate bycatches? p.175-184 in Nolan, C.P. (ed.), *Proceedings of the International Conference on Integrated Fisheries Monitoring.* Sydney, Australia, 1-5 February 1999.

Hay, D. E., R. Harbo, J. Boutillier, E. Wylie, L. Convey, and P.B. McCarter. 1999. Assessment of bycatch in the 1997 and 1998 shrimp trawl fisheries in British Columbia, with emphasis on eulachons. Canadian Stock Assessment Secretariat Document 99/179.

Hilborn, R. and Walters, C.J. 1992. *Quantitative fisheries stock assessment: choice, dynamics and uncertainty.* Chapman and Hall, New York.

Inter-American Tropical Tuna Commission. 2001. Annual Report of the Inter-American Tropical Tuna Commission, 2001.

Inter-American Tropical Tuna Commission. 2003. Agreement on the International Dolphin Conservation Program (amended June 2003). www.iattc.org

Karp, W. A. and H. McElderry. 1999. Catch monitoring by fisheries observers in the United States and Canada. p 261-284 in Nolan, C.P. (ed.), *Proceedings of the International Conference on Integrated Fisheries Monitoring.* Sydney, Australia, 1-5 February 1999.

Liggens, G. W., M. J. Bradley, S. J. Kennelly. 1997. Detection of bias in observer-based estimates of retained and discarded catches from a multispecies trawl fishery. *Fisheries Research* 32:133-147.

Medley, P. 2001. Estimating discards from catch species compositions. *Fisheries Center Research Report* 9 (3):46-52. University of British Columbia.

Mohn, R. 1999. The retrospective problem in sequential population analysis: an investigation using cod fishery and simulated data. *ICES Journal of Marine Science* 56:473-488.

National Marine Fisheries Service. 2000. Notice of Requirements of the Order of August 4, 2000, of the United States District Court for the District of Hawaii, 65 Fed. Reg. 49,968 (Aug. 16, 2000)

National Marine Fisheries Service 2001. Observer coverage plan: sampling plan and logistics for West Coast Groundfish Observer program. Fall 2001. Appendix A-statistical sampling design. Available from *http://www.nwfsc.noaa.gov/fram/Observer/ObserverSamplingPlan.pdf*

National Marine Fisheries Service. 2002a. Draft Environmental Assessment, Regulatory Impact Review and Initial Regulatory Flexibility Act Analysis (EA/RIR/IRFAA) Extending and Improving the North Pacific Groundfish Observer Program Beyond 2002.

National Marine Fisheries Service. 2002b. Atlantic highly migratory species; pelagic longline fishery; shark gillnet fishery; sea turtle and whale protection measures. *Federal Register* 67(131):45393-45401.

Nolan, C.P. (ed.), 1999. *Proceedings of the International Conference on Integrated Fisheries Monitoring.* Sydney, Australia, 1-5 February 1999.

Northwest Fisheries Science Center West Coast Groundfish Observer Program. 2003. Initial Data Report and Summary Analyses (http://www.nwfsc.noaa.gov/research/divisions/fram/observer/narjan03.pdf).

Ortiz, M., C. M. Legault and N. M. Ehrhardt. 2002. An alternative method for estimating bycatch from the U.S. shrimp trawl fishery in the Gulf of Mexico, 1972–1995. *Fishery Bulletin* 98:583-599.

Palka, D. L. and M. C. Rossman. 2001. Bycatch Estimates of Coastal Bottlenose Dolphin (*Tursiops truncatus*) in U.S. Mid-Atlantic Gillnet Fisheries for 1996 to 2000. Northeast Fisheries Science Center Reference Document 01-15.

Pikitch, E. K. and E. A. Babcock. 2002. Critique of the NMFS report, "Relative Precision of discard rate estimates for the Northeast groundfish complex," by M. J. Fogarty and W. L. Gabriel.

Pikitch, E. K., D.L. Erickson and J.R. Wallace.1988. An evaluation of the effectiveness of trip limits as a management tool. NWAFC Processed Report 88-27, 33 pp.

Pikitch, E. K. 1991. Technological interactions in the U.S. West Coast groundfish fishery and their implications for management. ICES Marine Science Symposium 193: 253-263.

Punt, A. 1999. Evaluating the costs and benefits of alternative monitoring programmes for fisheries management. p 209-222.in Nolan, C.P. (ed.), *Proceedings of the International Conference on Integrated Fisheries Monitoring.* Sydney, Australia, 1-5 February 1999.

Rao, P. S. R. S. 2000. *Sampling methodologies with applications.* Chapman and Hall/CRC. New York.

Rice, J. 1999. Stock assessments of target species. p. 51-64 in Nolan, C.P. (ed.), 1999. *Proceedings of the International Conference on Integrated Fisheries Monitoring.* Sydney, Australia, 1-5 February 1999

Rochet, M., I, Péronnet, and V. M. Trenkel. 2002. An analysis of discards from the French trawler fleet in the Celtic Sea. *ICES Journal of Marine Science* 59: 538–552.

Rohan, G. 1999. Ensuring monitoring contributes to the pursuit of management objectives: an Australian Fisheries Management Authority perspective. p 129-144 in Nolan, C.P. (ed.), 1999. *Proceedings of the International Conference on Integrated Fisheries Monitoring.* Sydney, Australia, 1-5 February 1999.

Saila, S. B. 1983. Importance and assessment of discards in commercial fisheries. FAO Fisheries Circular No. 765.

Sampson, D. 2002. Final Report to the Oregon Trawl Commission on Analysis of Data from the At-Sea Data Collection Project. Oregon State University. Newport, Oregon. *http://www.onid.orst.edu/~sampsond/projects/edcp/*

Smith, T. D. 1998. Simultaneous and complementary advances: mid-century expectations of the interaction of fisheries science and management. *Reviews in Fish Biology and Fisheries* 8:335-348.

Stratoudakis, Y., R. J. Fryer, R. M. Cook, and G. J. Pierce. 1999. Fish discarded from Scottish demersal vessels: Estimators of total discards and annual estimates for targeted gadoids. *ICES Journal of Marine Science* 56: 592–605.

Tamsett, D. and G. Janacek. 1999. Sampling trips for measuring discards in commercial fishing based on multilevel modeling of measurements in the North Sea from NE England. *Fisheries Research* 42:103-115

Venables, W. N., and B. D. Ripley. 1997. *Modern applied statistics with S-plus.* 2nd edition. Springer-Verlag, New York.

Walsh, W. A., P. Kleiber and M. McCracken. 2002. Comparison of logbook reports of incidental blue shark catch rates by Hawaii-based longline vessels to fishery observer data by application of a generalized linear model. *Fisheries Research* 58:79-94.

Williams, M. J. and V. P. Corral. 1999. Fisheries monitoring: management models, compliance and technical solutions. p 37-50 in Nolan, C.P. (ed.), *Proceedings of the International Conference on Integrated Fisheries Monitoring.* Sydney, Australia, 1-5 February 1999.

Tables

Table 1. Levels of coverage in NMFS observer programs. Levels of coverage are generally per unit of effort (e.g., fishing days, sets). From *http://www.st.nmfs.gov/st1/nop/*.

Observer Program	Level of coverage
Alaska Marine Mammal Observer Program	<5%
Offshore Pacific Whiting Fishery	100%
North Pacific and Bering Sea Groundfish Trawl and Fixed Gear Fishery Observer Program	Vessels >125ft.=100%, 60-124ft. =30%, Vessels<60ft.=0%
West Coast Groundfish Observer Program	Target 10%
Southeastern Shrimp Otter Trawl Fishery	<<1%
Hawaii Swordfish-Tuna Longline Observer Program	Level not specified
CA/OR Drift Gillnet Observer Program	~23%
West Coast Pelagic Longline Observer Program	Level not specified
SEFSC Pelagic Longline Observer Program	Target 5%
SER Shark Bottom Longline Observer Program	2-4%
SEFSC Shark Drift Gillnet Observer Program	100%
Northwest Atlantic Sustainable Fisheries Support	<1% in trawl fishery
New England and Mid-Atlantic Gillnet Fisheries	2-5%
Atlantic Sea Scallop Dredge Fishery - Georges Bank	25%

Table 2. Summaries of simulations based on U.S. Pacific and Atlantic gillnet fisheries. The total catch of the bycatch species was estimated with a combined ratio estimator that took into account correlations with the denominator variable. The final column indicates the fraction of vessel trips that had to be sampled to achieve the desired level of accuracy and precision. (See Appendix 1 and Figure 2 for details.)

Fishery	Bycatch sp.	Target	Denominator variable	Sampling fraction to get 90% within 10%
Pacific groundfish	Sablefish	Dover sole, sablefish, thornyheads	towing hours	30%-40%
Pacific groundfish	Dover sole	Dover sole, sablefish, thornyheads	towing hours	30%-40%
Pacific groundfish	Bocaccio	Flatfish	towing hours	>50%
Atlantic coastal gillnet	Bottlenose dolphin	Monkfish, striped bass, black drum, croaker, spiny dogfish	total catch of target species	>50%

Table 3. Summaries of simulated data applications for a generalized fishery. (See Appendix 1 and Figs. 3 and 4.)

(a) Effect of the rarity of the bycatch species on the required sampling fraction. The fisheries have 1000 trips per year, and differ only in the rarity of the bycatch species.

	Discard species increasingly common in the catch ?			
Total bycatch as percent of total catch plus bycatch	0.1	0.7	6	35.4
Percent coverage to get within 10% of the correct value in at least 90% of simulations	>50	28	18	17

(b) Comparison between sampling fraction and sample size. The fisheries vary only in the number of trips. Smaller fisheries require a higher sampling fraction but a lower sample size.

Number of trips in fishery	10000	1000	100
Percent coverage to get within 10% of the correct value in at least 90% of simulations	3.6	28	>50
Sample size to get 90% within 10%	360	280	50-100

Figures

Figure 1. Average sampling error for 500 simulated observer samples at varying levels of observer coverage, for a large fishery and a small fishery. For the small fishery (100 trips), the smallest stratum was not adequately covered at low levels of sampling coverage, and there was a strong negative bias in the total discard estimates.

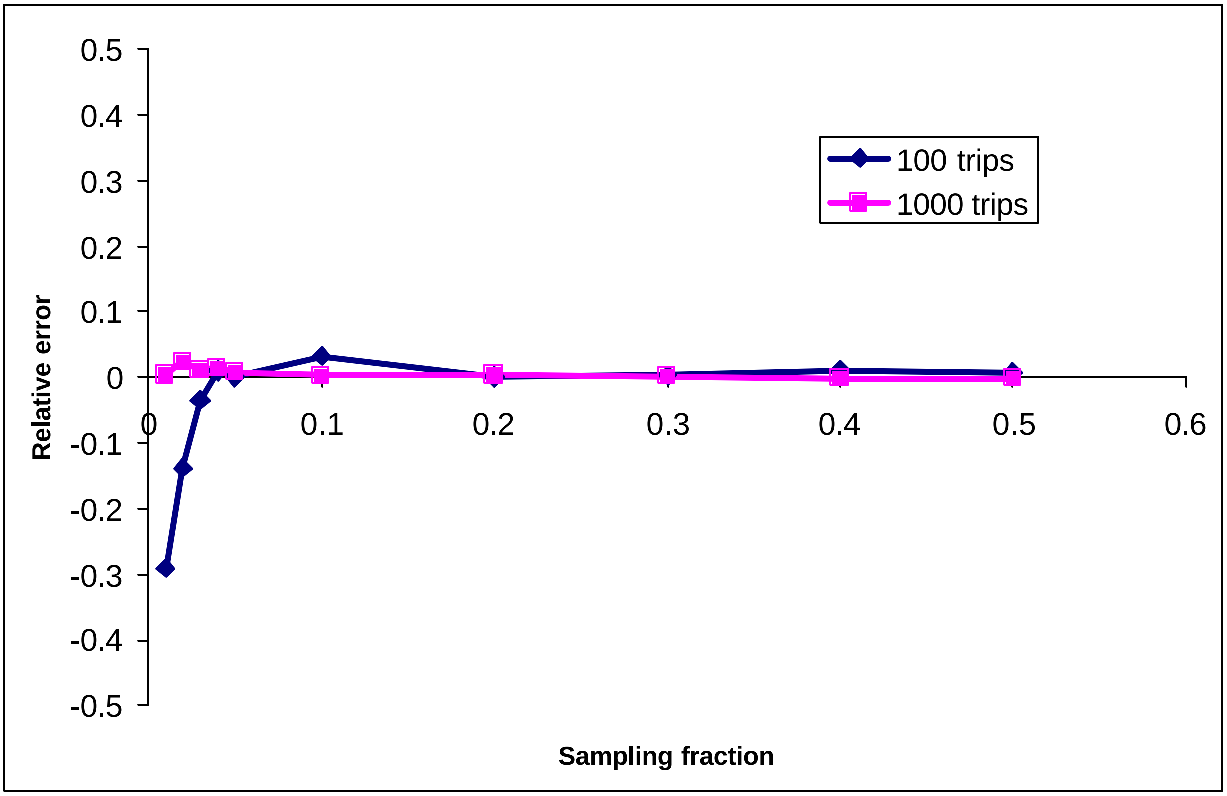


Figure 2. Percent of 500 simulated observer samples that gave an estimate of total bycatch within 10% of the true value, for simulated fisheries based on (a) bycatch of Dover sole or sablefish in the U.S. Pacific groundfish fishery targeting the Dover sole-thornyhead-sablefish complex, (b) bycatch of bocaccio rockfish in the U.S. Pacific groundfish fishery targeting flatfish, and (c) bycatch of bottlenose dolphins in a mixed-species Atlantic coast gillnet fishery. (See Appendix 1 and Table 2.)

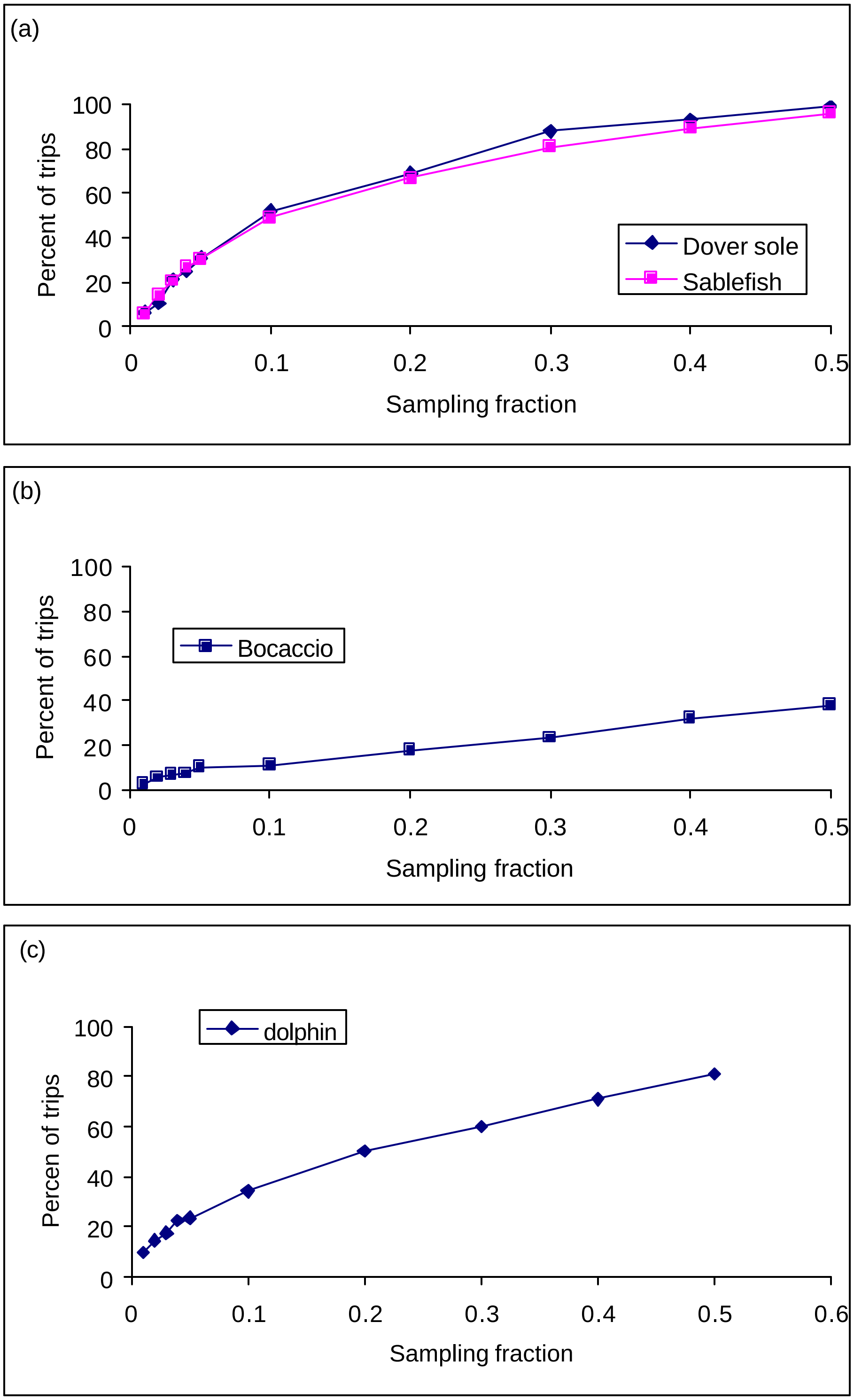


Figure 3. Percent of 500 simulated observer samples that gave an estimate of total bycatch within 10% of the true value, for discard species varying from rare to common in a hypothetical fishery where only the rarity of the discard species varied. (See Appendix 1 and Table 3.)

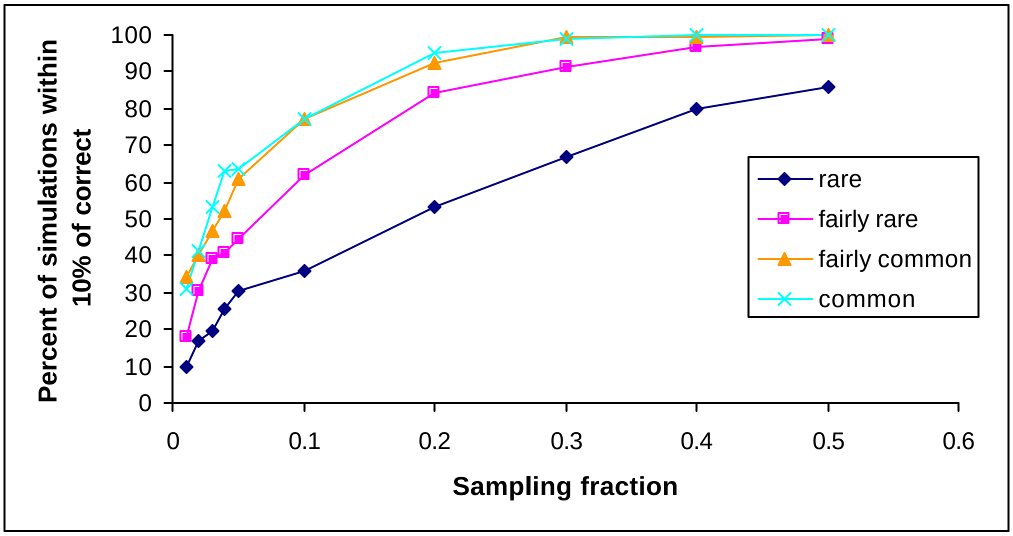
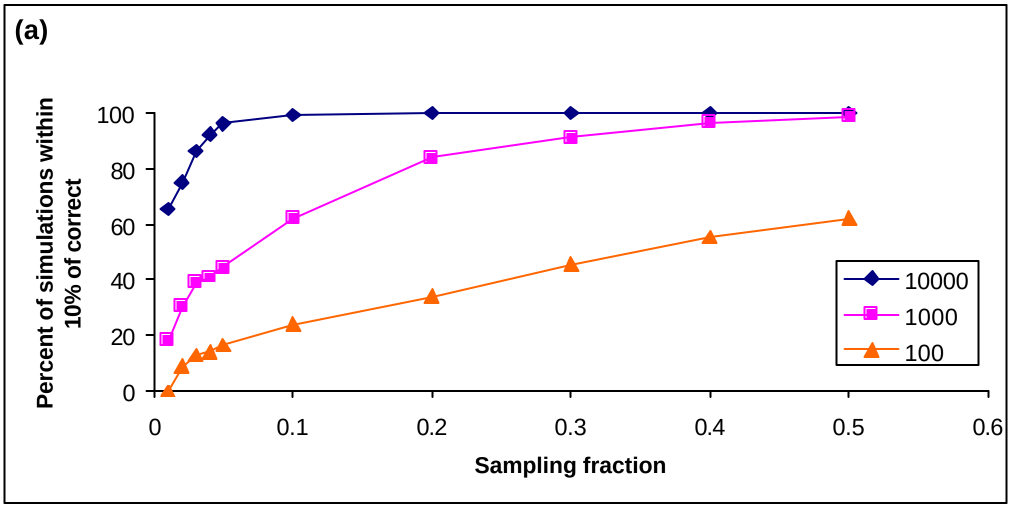
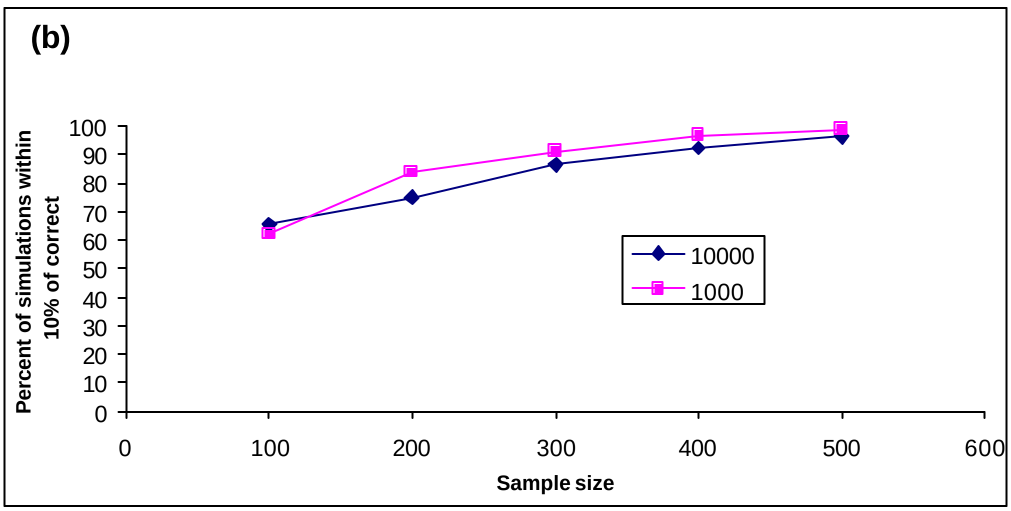
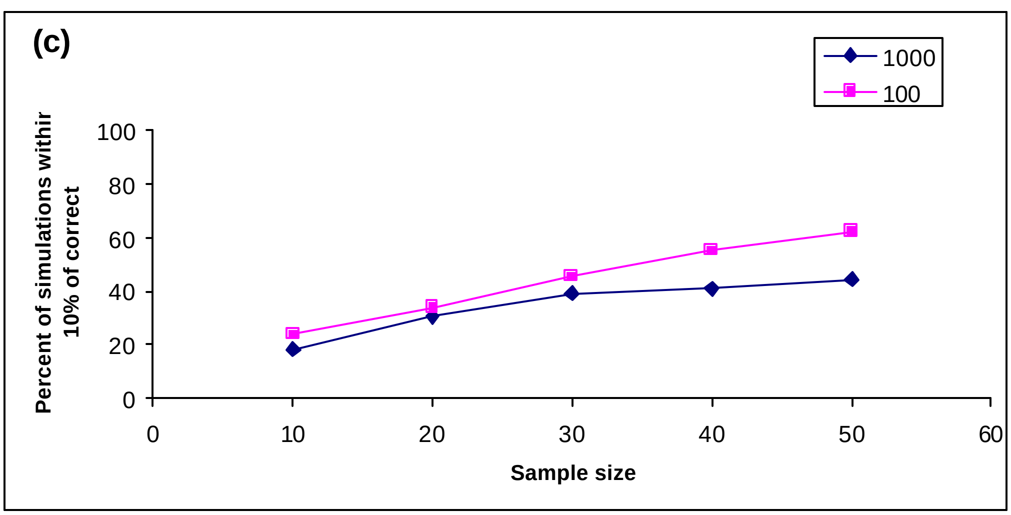


Figure 4. Percent of 500 simulated observer samples that gave an estimate of total bycatch within 10% of the true value, with (a) increasing sampling fraction and (b) and (c) increasing sample size, for hypothetical simulated fisheries varying only in the number of trips (shown in the legend). (See Appendix 1 and Table 3.)







Appendix 1. Methodology used for simulated data applications.

The objective of this simulation exercise was not to exhaustively cover all possible kinds of data that might be generated by an observer sampling program. Rather, the objective was to explore some common characteristics of observer data and determine general rules of thumb for assigning levels of observer coverage. The simulated data application was developed using S-plus 4.5 for Windows. All random numbers were drawn using the S-plus pseudo-random number generators (*rnorm, rlognorm*, etc.) with the default random number seeding (Venables and Ripley 1997).

We simulated two kinds of fisheries. The first, a generalized hypothetical fishery, was used to examine the effect of species rarity and other factors on the required level of coverage, by holding all other factors constant across fisheries. The second type of fishery was a more realistic simulation based on several actual fisheries.

For each simulation exercise, the following algorithm was used:
1. Simulate a fishery (the sampling universe)
2. Simulate the observer sampling process, and repeat many times for various levels of observer coverage
3. Estimate the total bycatch for each sample at each level of coverage
4. Compare the estimated total bycatch at each coverage level to the "true" values from the simulated fishery
5. Repeat for different simulated fisheries

The characteristics of each simulated fishery are described below. For each simulated fishery, the observer program was simulated by randomly drawing trips from the universe of trips, assuming that the observer would sample all sets for the trips they observed. The summary statistics generated were the estimate of total discard from a pooled ratio estimator (Cochran 1977):

$$(A1)\quad \hat{D}_T = \frac{\sum_i D_i}{\sum_i C_i} C_T$$

where $\hat{D}_T$ is the estimated total bycatch, and C_i and D_i are the catch and bycatch, respectively, in observer sample i, and C_T is the total catch from the entire fishery, assumed known without error from landings data.

For some simulations, we simulated a fishery that had different discard rates in different strata (areas or gears), with the trips randomly allocated to strata. In this case, the total discard in each stratum was estimated with equation A1, and the estimates were added to get the total for the fishery.

We did 500 Monte Carlo simulations at each of 10 levels of observer coverage, and summarized the simulations with total relative error, calculated as:

$$\text{(A2)}\; error = \frac{\hat{D}_T - D_T}{D_T}$$

where D_T is the known total bycatch from the fishery. This measure of error includes both accuracy and precision as it is calculated relative to the actual known value of total bycatch instead of the estimated mean. At each level of observer coverage, the fraction of simulated observer estimates of bycatch for which the error is less than 10% was graphed, and the fraction required to achieve 90% of the simulations within 10% was calculated.

Generalized hypothetical fisheries

The generalized hypothetical fisheries were simulated as a set of independent fishing trips, each of which consisted of a number of (possibly correlated) sets. The number of sets in each trip was randomly drawn from a binomial distribution. In some cases all trips were considered equivalent, in some cases they were stratified (e.g. by area).

For all simulated fisheries, there were assumed to be two species caught—the target catch species and the discarded bycatch species. The distribution of catches $C_{t,s,i}$ of the target species in trip t, strata s and set i was drawn from a Delta-lognormal distribution with a probability of a positive catch in a trip equal to $p_{c,t,s}$, and the distribution of the catches in positive trips was drawn from a lognormal distribution with the log mean equal to $\overline{C}_{t,s}$ and the log standard deviation equal to s_C.

The probability of a positive catch in each trip $p_{c,t,s}$ and stratum and the mean catch in each trip and stratum $\overline{C}_{t,s}$ were drawn from linear models that included a stratum effect and a trip effect, each drawn from a normal distribution:

$$(A3)\quad \overline{C}_{t,s}, = m_{c,t} + m_{c,s} \quad \text{and}$$

$$(A4)\quad p_{c,t,s}, = r_{c,t} + r_{c,s}$$

The bycatch in each $D_{t,s,i}$ set was calculated similarly from a Delta-lognormal distribution, with a probability of positive bycatch and mean positive bycatch drawn from a linear model, except thatthere was also a linear effect of catches in a set on bycatch in the same set:

$$(A5)\quad \overline{D}_{t,s,i}, = m_{d,t} + m_{d,s} + m_{d,c} \cdot C_{t,s,i} \quad \text{and}$$

$$(A6)\quad p_{d,t,s,i}, = r_{d,t} + r_{d,s} + r_{d,c} \cdot C_{t,s,i}$$

See Table A1 and A2 for definitions of the parameters and their values for each simulated fishery. The issues addressed were stratification, number of trips in the fishery, and the rarity of the bycatch species.

Pacific coast groundfish trawl fishery simulator

To examine the accuracy and precision of the estimates of total bycatch in the trawl fishery off the coasts of Washington, Oregon, and California, we simulated a fishery as similar as possible to the actual fishery in the years 2001 and 2002 (Table A3, Table A4). We simulated the catches of Dover sole and sablefish, which are two common species in the DTS (Dover sole, sablefish and thornyhead) complex North of 40°10'N, and a rare species (bocaccio) in the flatfish complex all along the coast. The mean and standard error of the mean of sablefish and Dover sole catch per hour in each stratum for the DTS complex in the North area were taken from Northwest Fisheries Science Center (2003), Table 4. The variance in catch per hour was assumed to be equal to the standard error of the mean catch per hour squared times the sample size. The distribution of catch per hour per tow was assumed to be lognormal for Dover sole and sablefish, with log-mean and log-standard deviations calculated from the mean and variance in each stratum.

For bocaccio caught as bycatch in the flatfish fishery, because the species is not caught in most tows, the catches were drawn from a Delta-lognormal distribution. The probability of a positive catch was assumed to be 0.019 in the North and 0.094 in the South, based on Fig. 6 in NWFSC (2003). The mean and variance of catches for positive trips in each stratum were calculated as the mean and variance calculated as above from NWFSC (2003) Table 4, divided by the probability of a positive trip (for the mean) and the probability of a positive trip squared (for the variance).

The number of trips was assumed to be 1000 for the DTS complex in the North, and 3500 for the coastwide flatfish fishery. The distribution of tows in each trip and the duration of each tow were calculated based on observer data from 1989-1990 (Babcock and Pikitch 2000), as this information has not been published from the current observer study. The number of DTS tows in each trip was assumed to be binomial with a mean of 8 and maximum number of tows of 20, and the duration of each tow was assumed to be Poisson with the mean shown in Table A4. The number of flatfish tows in each trip was assumed to binomial with a mean of 4 and maximum of 20, and the tow duration was Poisson with the lambda shown in the table calculated from Babcock and Pikitch (2000).

Atlantic coast bottlenose dolphin bycatch in gillnet fishery simulator

Bottlenose dolphins are caught very rarely in the Atlantic gillnet fishery, thus dolphin catch can be modeled as a Poisson process (Palka and Rossman 2001). The catch of the target species per set was assumed to be lognormal with a mean equal to the total observed catch divided by the number of observed sets in each stratum in the year 2000, and a 200% CV, corresponding to the log-mean and log-sd shown in Table A5. The ratio of dolphin bycatch to target species catch was modeled as a Poisson process with a mean equal to the bycatch rate in each strata in 2000 calculated by Palka and Rossman (2001). Trips were allocated to management unit and area strata by the proportion of observed trips in each category in the year 2000, and were further allocated to mesh size strata based on the fraction of catch taken by each mesh size in each area x management unit. The number of trips was assumed to be 15,000, with the number of sets in each trip drawn from a binomial distribution with mean 3.2, and maximum number of sets 10 (Table A4).

Although the total bycatch was estimated in this study using a generalized linear model (GLM) estimator (Ortiz et al. 2000) instead of a ratio estimator, we estimated total bycatch with a ratio estimator based on the ratio of bycatch to total catch.

Table A1. Input parameter values for the default generalized fishery simulation.

Parameter	Definition	Distribution	Number of values drawn	Central parameter	Shape parameter
T	number of trips	constant	1	1000	
N_t	sets per trip	binomial	T	p=0.1	n=100
A	number of strata	constant	1	1	
$r_{C,t}$	catch probability trip effect	normal	T	0.90	0.02
$r_{C,s}$	catch probability stratum effect	normal	A	0	0.1
$m_{C,t}$	positive catches trip effect	normal	T	50	0.1
$m_{C,s}$	positive catches stratum effect	normal	A	0	0.1
$r_{D,t}$	bycatch probability trip effect	normal	T	0.1	0.01
$r_{D,s}$	bycatch probability stratum effect	normal	A	0	0.01
$r_{D,c}$	positive bycatch catch effect	normal	$\sum N_t$	0.00001	0.00001)
$m_{D,t}$	positive bycatch trip effect	normal	T	5	0.01
$m_{D,s}$	positive bycatch stratum effect	normal	A	0	0.01
$m_{D,c}$	positive bycatch catch effect	normal	A	0.001	0.001
s_C	log standard deviation of the catches in positive sets	constant	1	2.0	
s_D	log standard deviation of the bycatch in positive sets	constant	1	0.5	

Table A2. Description of the simulation experiments. All simulated fisheries were as defined for the default fishery (Table A1), except where noted.

Description	Parameter values
Comparing large and small fisheries with a stratified estimator	Trips were allocated to three strata with probability (0.1,0.25,0.65), $r_{C,s}$ was Normal(0,0.001), $m_{C,s}$ was Normal(0,0.1), $r_{D,s}$ was Normal(0,.001), $m_{D,s}$ was Normal(0,0.01). Number of trips (T) was either 100 or 1000 (Fig. 1)
Bycatch species rarity varying	Rare: $r_{D,c}$ was Normal(0.000001,0.000001), $r_{D,t}$ was Normal(0.03,0.001), $m_{D,t}$ was Normal(2,0.1), $m_{D,c}$ was Normal(0.000001,0.000001) Fairly rare: Default fishery Fairly common: $r_{D,c}$ was Normal(0.0001,0.0001), $r_{D,t}$ was Normal(0.2,0.01), $m_{D,t}$ was Normal(20,0.05) Common: $r_{D,c}$ was Normal(0.00001,0.0001), $r_{D,t}$ was Normal(0.5,0.02), $m_{D,t}$ is Normal(50,0.1), s_D=1.0 (Fig. 3, Table 3a)
Number of trips varying	Default fishery, except that number of trips (T) was 100, 1000 or 10,000 (Fig. 4, Table 3b)

Table A3. Inputs for the simulations based on the Pacific observer program, for the DTS strategy North of 40°10’N (NWFSC 2003, tow duration from Babcock and Pikitch 2000).

					Sablefish per hour				Dover sole per hour			
Stratum	Depth (fm)	Period	Fraction of tows	Tow mean(hr)	mean	sd	log-mean	log-sd	mean	sd	log-mean	log-sd
1	0-100	Sep-Oct 2001	0.04	2.5	34.59	82.73	2.59	1.38	52.41	109.47	3.12	1.30
2	0-100	Jan-Feb 2002	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3	0-100	Mar-Apr 2002	0.04	2.5	85.16	290.45	3.18	1.59	59.51	75.65	3.61	0.98
4	0-100	May-Jun 2002	0.12	2.5	179.10	1047.7	3.41	1.89	20.54	33.74	2.37	1.14
5	0-100	Jul-Aug 2002	0.06	2.5	163.80	368.32	4.20	1.34	32.10	51.49	2.83	1.13
6	100-200	Sep-Oct 2001	0.04	2.9	14.96	25.78	2.02	1.17	18.12	34.88	2.12	1.24
7	100-200	Jan-Feb 2002	0.01	2.9	146.46	185.38	4.51	0.98	22.26	27.11	2.65	0.95
8	100-200	Mar-Apr 2002	0.04	2.9	103.55	189.87	3.90	1.21	79.28	321.57	2.94	1.69
9	100-200	May-Jun 2002	0.03	2.9	51.98	105.81	3.13	1.28	20.67	39.21	2.27	1.24
10	100-200	Jul-Aug 2002	0.03	2.9	48.21	94.76	3.08	1.26	36.21	80.76	2.70	1.34
11	>200	Sep-Oct 2001	0.08	5.4	2.11	5.88	-0.34	1.47	30.88	42.36	2.90	1.03
12	>200	Jan-Feb 2002	0.18	5.4	16.33	42.07	1.78	1.43	9.03	31.36	0.92	1.60
13	>200	Mar-Apr 2002	0.26	5.4	20.66	80.39	1.64	1.67	11.96	39.74	1.24	1.58
14	>200	May-Jun 2002	0.06	5.4	21.13	41.06	2.27	1.25	36.93	70.79	2.84	1.24
15	>200	Jul-Aug 2002	0.01	5.4	22.44	31.75	2.56	1.05	13.99	37.85	1.58	1.46

Table A4. Inputs for bocaccio bycatch in the Pacific coast flatfish fishery (NWFSC 2003, tow duration from Babcock and Pikitch 2000). Areas are North and South of 40°10'N, Oregon.

Stratum	Depth (fm)	Period	Area	Frac-tion of tows	Prob. of positive set	Mean set (hrs)	Bocaccio catch per hour mean	variance	log-mean	log-sd
1	<100	Sep-Oct 2001	N	0.071	0.019	2.4	0.065	0.151145	-0.591	1.901
2	<100	Nov-Dec 2001	N	0.043	0.019	2.4	0.009	0.007216	0	0.001
3	<100	Jan-Feb 2002	N	0.01	0.019	2.4	0.251	1.262893	1.05	1.745
4	<100	Mar-Apr 2002	N	0.1	0.019	2.4	1.825	96.61542	2.855	1.844
5	<100	May-Jun 2002	N	0.224	0.019	2.4	0.061	0.291242	-1.032	2.092
6	<100	Jul-Aug 2002	N	0.257	0.019	2.4	1.596	326.5109	1.991	2.205
7	100-200	Sep-Oct 2001	N	0.012	0.019	3.3	0.000	36385.47	0	0.001
8	100-200	Nov-Dec 2001	N	0.014	0.019	3.3	0.000	6123.982	0	0.001
9	100-200	Jan-Feb 2002	N	0.039	0.019	3.3	0.502	4.002719	1.85	1.681
10	100-200	Mar-Apr 2002	N	0.017	0.019	3.3	0.075	0.187286	-0.395	1.878
11	100-200	May-Jun 2002	N	0.002	0.019	3.3	0.000	20171.56	0	0.001
12	100-200	Jul-Aug 2002	N	0.011	0.019	3.3	0.000	78868.06	0	0.001
13	<200	Jan-Feb 2002	N	0.025	0.019	4.1	0.088	0.264889	-0.266	1.889
14	<200	Mar-Apr 2002	N	0.012	0.019	4.1	0.000	17212.88	0	0.001
15	<200	Jul-Aug 2002	N	0.001	0.019	4.1	0.000	0	0	0.001
16	<100	Sep-Oct 2001	S	0.066	0.094	4.1	3.206	242.6569	1.929	1.79
17	<100	Nov-Dec 2001	S	0.018	0.094	4.1	24.621	4783.465	4.476	1.478
18	<100	Jan-Feb 2002	S	0.028	0.094	4.1	1.378	8.266334	1.847	1.295
19	<100	Mar-Apr 2002	S	0.019	0.094	4.1	0.748	3.468673	1.087	1.405
20	<100	May-Jun 2002	S	0.004	0.094	4.1	3.909	61.06034	2.924	1.268
21	100-200	Sep-Oct 2001	S	0.018	0.094	4.1	2.889	33.85954	2.616	1.273
22	100-200	Nov-Dec 2001	S	0.003	0.094	4.1	11.868	470.1041	4.106	1.211
23	100-200	Jan-Feb 2002	S	0.003	0.094	4.1	1.931	15.35548	2.207	1.278
24	100-200	Mar-Apr 2002	S	0.001	0.094	4.1	0.000	0	0	0.001
25	100-200	Jul-Aug 2002	S	0.002	0.094	4.1	0.612	1.123632	1.181	1.177
26	<200	Sep-Oct 2001	S	0.002	0.094	4.1	0.000	256.6836	0	0.001

Table A5. Inputs for the simulations based on the Atlantic gillnet fishery observer program for bottlenose dolphin bycatch (Palka and Rossman 2001).

Strata	Mesh Size	Management unit	Area	Bycatch rate	Fraction of sets	Log-mean	Log-sd
1	Small	Winter - VA Mixed Stock	state	0.0075	0.084	-3.241	1.269
2	Small	Winter - VA Mixed Stock	federal	0.0007	0.003	-2.843	1.269
3	Small	Summer - Northern Migratory	state	0.0266	0.078	-2.956	1.269
4	Small	Summer - Northern Migratory	federal	0.0024	0.03	-2.979	1.269
5	Small	Winter - North Carolina Mixed Stock	state	0.0159	0.207	-3.702	1.269
6	Small	Winter - North Carolina Mixed Stock	federal	0.0014	0.05	-2.875	1.269
7	Small	Summer - Northern North Carolina	state	0.0698	0.076	-3.765	1.269
8	Small	Summer - Northern North Carolina	federal	0.0066	0.002	-4.039	1.269
9	Small	Summer- Southern North Carolina	state	0.0006	0.037	-2.945	1.269
10	Small	Summer- Southern North Carolina	federal	0.0001	0.001	-5.156	1.269
11	Medium	Winter - VA Mixed Stock	state	0.0243	0.051	-1.318	1.269
12	Medium	Winter - VA Mixed Stock	federal	0.0022	0.019	-1.709	1.269
13	Medium	Summer - Northern Migratory	state	0.0824	0.005	-2.828	1.269
14	Medium	Summer - Northern Migratory	federal	0.0079	0.011	-2.961	1.269
15	Medium	Winter - North Carolina Mixed Stock	state	0.0504	0.153	-2.029	1.269
16	Medium	Winter - North Carolina Mixed Stock	federal	0.0047	0.052	-1.5	1.269
17	Medium	Summer - Northern North Carolina	state	0.1979	0.019	-4.398	1.269
18	Medium	Summer - Northern North Carolina	federal	0.0213	0.001	-3.417	1.269
19	Medium	Summer- Southern North Carolina	state	0.002	0	-6.95	1.269
20	Medium	Summer- Southern North Carolina	federal	0.0002	0	0	1
21	Large	Winter - VA Mixed Stock	state	0.1825	0.067	-3.563	1.269
22	Large	Winter - VA Mixed Stock	federal	0.0193	0.006	-2.841	1.269
23	Large	Summer - Northern Migratory	state	0.4458	0.002	-6.103	1.269
24	Large	Summer - Northern Migratory	federal	0.0663	0.017	-5.41	1.269
25	Large	Winter - North Carolina Mixed Stock	state	0.3222	0.005	0	1
26	Large	Winter - North Carolina Mixed Stock	federal	0.0403	0.024	-2.429	1.269
27	Large	Summer - Northern North Carolina	state	0.6885	0	0	1
28	Large	Summer - Northern North Carolina	federal	0.1632	0	-6.103	1.269
29	Large	Summer- Southern North Carolina	state	0.0173	0	0	1
30	Large	Summer- Southern North Carolina	federal	0.0016	0	0	1