Issue Paper: The Application of Basket Sampling Methods to Estimate Bycatch Rates of Prohibited Species Aboard Domestic Fishing Vessels

By

Gordon H. Kruse

Regional Information Report[1] No. 5J91-06
Alaska Department of Fish & Game
Division of Commercial Fisheries
P.O. Box 3-2000
Juneau, Alaska 99802-2000

March 18, 1991

[1]The Regional Information Report Series was established in 1987 to provide an information access system for all unpublished divisional reports. These reports frequently serve diverse ad hoc informational purposes or archive basic uninterpreted data. To accommodate timely reporting of recently collected information, reports in this series undergo only limited internal review and may contain preliminary data; this information may be subsequently finalized and published in the formal literature. Consequently, these reports should not be cited without prior approval of the author or the Division of Commercial Fisheries.

Issue Paper: The Application of Basket Sampling Methods to Estimate Bycatch Rates of Prohibited Species Aboard Domestic Fishing Vessels

by
Alaska Department of Fish and Game
Division of Commercial Fisheries

Introduction

New sampling procedures are being proposed for the estimation of bycatch rates by domestic fishing vessels. These new proposed protocols are described in section 4, "Observer sampling procedures," in a draft text to be submitted to the *Federal Register* (NMFS 1991). In overview, the plan is to sample at least 50% of vessel's hauls while an observer is aboard. The hauls to be sampled will be predetermined using a random number table. From each haul, the observer will collect samples for estimation of species composition and prohibited species bycatch rates for each vessel on a monthly basis. A minimum of three basket samples (a total of 100 kg) will be sampled from each haul. The samples are to be taken from one of the following four locations: (1) unsorted codend of the trawl as it is emptied; (2) unsorted catch as it falls from a holding bin onto a conveyor belt; (3) unsorted catch dumped into a holding bin; or (4) unsorted fish as the trawl is dumped and spread onto the deck. The contents of each sample basket will be enumerated and weighed.

Concerns with Basket Sampling Approach

This proposed new basket sampling procedure for observers raises concerns about the accuracy of resultant bycatch estimates, because it has been well established that it is extremely difficult to collect representative samples using a basket sampling technique on unsorted trawl catches (Pope 1963; Paloheimo and Dickie 1963; and May and Hodder 1966). Basket samples are heterogenous with respect to fish length and species composition (Paloheimo and Dickie 1963). In fact, based upon their field studies, Paloheimo and Dickie (1963) concluded that *"In general it appears that the basket-lot technique of handling individual species of fish is not a satisfactory procedure on which to base a sampling program. Heterogeneity among the baskets is so great that random sampling for lengths is impossible, and sampling for counts requires handling of a high proportion of the catch."* Pope (1963) agreed that *"Any attempt to select representative samples before sorting is ... likely to produce substantial bias."*

Several factors adversely affect the ability to obtain representative samples for the calculation of accurate, precise estimates of bycatch rates. First, it is well known that fish stratify by size and species within a trawl. Because of this stratification, samples from no one section of the trawl can be considered "representative." Second, it is

difficult to collect random samples within any one section of a haul, because there is a tendency for the smaller fish to lie on the bottom of the pile and the larger fish to lay on top or to slide to the edges (Paloheimo and Dickie 1963; May and Hodder 1966).

These problems may persist in the case of bycatch in groundfish fisheries off Alaska. In instances of extremely high crab bycatch, large concentrations of red king crab, called "pods," may be caught. When captured by the net, stratification by species is extreme; a pod may remain intact and occur in only one portion of the trawl. Even in trawl catches with few crabs, all crabs would enter the entrance to the trawl along the bottom of the net. Because their body shape is very different from any fish species, differential hydrodynamics may lead to stratification in the trawl. Because halibut reach sizes much larger than most other fishes, they may stratify in the net, as well. When the trawl contents are dumped on deck, there may be a tendency for large crabs and halibut to lay on top of the pile due to their size. Further, any tendencies for crab to walk and halibut to flop off the pile toward the periphery of the work area will confound the ability of an observer to collect representative samples. There is the remaining problem of sampling a large halibut that, quite simply, will not fit into a sampling basket.

Problems in collecting representative samples will affect the accuracy of bycatch estimates derived from basket samples of domestic catches. Because it is difficult to develop unbiased estimates of species composition and fish size from basket samples even for the common species in the catch (Paloheimo and Dickie 1963), it is highly unlikely that a basket sampling plan would yield accurate estimates of bycatch rates of "rare" prohibited species in terms of weight of bycatch species per unit weight of all species in the haul. Indeed, Cochran (1977) pointed out that bias is of the order of $1/n$. In the case of three basket samples from one haul, the average bycatch rate per haul would be biased by 33%. Because of the possibility of systematic bias among all samples, it is not obvious that this sampling procedure will result in accurate estimates of monthly bycatch rates for each vessel.

An Example of Sampling Difficulties

As an example, suppose that an individual trawl haul caught 10 mt (10,000 kg) of fish and shellfish. Also, suppose that this haul caught 5 halibut weighing 10 kg each. The true bycatch rate of halibut is 5 halibut per 10 mt (0.5/mt) or 50 kg/10 mt (5 kg/mt). In terms of absolute weight, the halibut comprise 0.5% of the total catch; i.e., halibut is a 'rare' species in the catch. Last, suppose that three basket samples (a total of 100 kg) are taken from this haul.

There are a limited number of sampling outcomes for halibut. Conceivably, anywhere from none to all of these halibut could be sampled. In terms of weight, there are five possible outcomes for bycatch estimates: 0 kg/mt, 100 kg/mt, 200 kg/mt, 300 kg/mt,

400 kg/mt, and 500 kg/mt. Given that the true rate is 5 kg/mt, neither outcome can be considered as an accurate estimator of halibut bycatch in this haul.

If no halibut are subsampled, then the number of halibut in the three individual subsamples will be 0, 0, and 0. The mean is 0 kg/mt and the variance is 0 kg/mt. Thus, despite the high level of precision in this sample, the 95% confidence interval for this sample does not include the true value and the mean is biased. On the other hand, if only one halibut occurs in one subsample, the three observations are 0, 0, and 300 kg/mt. The mean will be 100 kg/mt and the variance will be 30,000 kg/mt! In such case the mean is biased high by a factor of 20, and the estimate is very imprecise. In the chance occurrence that all five halibut occur in the subsamples the estimated bycatch rate will be two orders of magnitude higher than the true rate, and the 95% confidence interval may not include the true rate.

This example demonstrates that representative samples of halibut bycatch rate for this haul are not possible with a basket sampling technique. All outcomes are either very precise but inaccurate, or very imprecise and inaccurate. No outcomes are accurate, regardless of precision. That is, bias occurs regardless of whether stratification of fish by size and species occurs when caught by a trawl or when piled on a deck. It is difficult to extend these results to, say, 30 such hauls sampled in one month aboard one vessel. If bycatch species were randomly distributed among all fish in the trawl and if observers truly select random samples, then the mean of 30 bycatch estimates may be unbiased. However, if observer-based estimators of size frequencies and species compositions of individual trawl hauls is consistently biased in the same direction, then the monthly mean bycatch estimate of each vessel will be biased.

Collecting Representative Basket Samples

In the observer sampling plan it is acknowledged that stratification by size or species may occur within the trawl, and it is mentioned that the observer will sample from different parts of the haul. However, Pope (1963) cautioned that *"in particular it could be extremely dangerous to 'select' a 'representative' basket as the sample."* The findings and conclusions of Pope (1963), Paloheimo and Dickie (1963), and May and Hodder (1966) suggest that no basket sampling plan applied to unsorted fish catches can produce unbiased estimates of catches. Furthermore, the potential problem of stratification of fish by size within a pile is not addressed by the new sampling plan.

Paloheimo and Dickie (1963) and Pope (1963) suggested that some special treatment is required to obtain representative basket samples of the catch of research vessels. These authors recommended first sorting the catch by species. Second, to collect representative basket samples, Paloheimo and Dickie (1963) suggested mixing basket lots of the catch pre-sorted by species and subsampling from these mixtures. Further, Pope (1963) suggested sampling the top, middle, and bottom of each basket.

An alternative method was described by May and Hodder (1966) in which the catch was sorted one species at a time into multiple baskets, and one or more baskets were selected at random for length frequency sampling as they were filled. Their method would not be applicable to determinations of species composition.

It does not appear that the recommended modifications to basket sampling suggested by Paloheimo and Dickie (1963), Pope (1963) and May and Hodder (1966) can be implemented on commercial vessels due to time constraints. If implemented, sorting by species, mixing of samples, and subsampling within baskets would cause undue costs on commercial operations. It is generally impractical to sort the entire catch by species aboard a commercial vessel, and it would be infeasible to create mixtures of baskets within species groups in any case.

Legal Aspects of Bycatch Rates

Aside from questions of accuracy of estimates based on a basket sampling program, the use of a 95% confidence interval for court proceedings may be questionable. The statistical power of this test has not been determined. Yet, it is not inconceivable that the proposed sampling plan would result in 95% confidence intervals that generally include zero.

Even if precise estimates of monthly bycatch rates were obtained, difficulties may arise in court cases. The rejection of the null hypothesis, when it is in fact true, is called a *type I error*. In a hypothesis using 95% confidence levels, 5% of the rejections of the null hypothesis (H_0) will be incorrect. Because court cases are likely to be made on an individual basis, it may be very difficult to demonstrate that the plaintiffs are not making a type I error, and that the defendant is not one of the 5% whose sample rate was statistically higher than the mean when his true rate was actually less than or equal to the mean.

Furthermore, the findings of Paloheimo and Dickie (1963), Pope (1963) and May and Hodder (1966) should be sufficient evidence to bring the validity of bycatch estimates obtained by basket samples into question in a court of law. In any case, the lack of data analysis to support the use of basket sampling procedures prohibit a rebuttal of findings from these earlier studies.

Alternative Estimators for the Incentive Program

It is generally true that the number of species in a sample increases asymptotically as sample size increases. The last species included in the sample are rare species. Because prohibited species are generally "rare," it follows that a large sample size is required to enumerate bycatch. Stated another way, *"if the number of individuals bearing a specified characteristic (such as having a length in a specified size range or being infected with a parasite) is to be estimated, the sample size necessary to ensure*

a maximum error not greater than some specified fixed amount increases rapidly as the true proportion bearing this characteristic decreases" (Pope 1963).

Because bycatch rates are small proportions, large sample sizes are required to obtain accurate estimates. In general, simple random sampling is an expensive method of estimating the number of rare species within an acceptable coefficient of variation (Cochran 1977). Rather, a whole haul sampling plan may be most appropriate method for estimation of bycatch rates for prohibited species. As Pope (1963) pointed out, *"if interest is confined to the total number of fish of a given species in a catch, it will often be as easy to make a complete count as to estimate this number from a sample."* If whole haul sampling is not feasible on all vessels due to size of catch or quantity of bycatch, the plan could be modified such that a fraction of the whole haul could be sampled. Estimates of bycatch derived from the sampled catch could be used for the incentive program. For monitoring bycatch caps, sampling results (e.g., bycatch rate) from the fraction sampled could be expanded to the entire haul by the estimate of total catch weight.

Whole haul sampling methods were considered but rejected in the observer plan, because the *"the nature and magnitude of the error associated with total catch estimates are unknown and cannot be measured."* The sampling plan goes on to state that without this information, statistical relationships between a vessel's bycatch rate and the bycatch standard cannot be determined. The apparent need for a confidence interval for court proceedings led to the development of the basket sampling proposal.

There appears to be a *"catch 22."* Confidence intervals cannot be developed for the most accurate estimator of bycatch (whole haul counts), yet confidence intervals can be calculated for a inaccurate estimator (basket sampling method). There are at least two solutions: (1) require scales on each vessel so that the total haul weight and its error structure is known; or (2) define "dirtiness" using an estimator other than a bycatch rate based upon total haul (or sample) weight of all species.

"Dirtiness" could be defined in terms of bycatch rate in which weight (or numbers) of prohibited species constitutes the numerator and some quantity other than total weight constitutes the denominator. The denominator could be a single haul, or a unit of time (e.g., hour), or some other easy-to-measure parameter with known error structure. If the entire haul is sampled and an error-free parameter (e.g., per haul) is chosen for the denominator, then the bycatch rate is known without error. With bycatch rate defined in this way, it is possible to calculate 95% confidence intervals for the monthly mean bycatch rate for each vessel.

Alternatively, "dirtiness" could be defined, not as a bycatch rate, but as a bycatch level. A vessel could be subject to penalty, if, for example it had more than x hauls or x percent of hauls per month in which bycatch exceed y crab or z halibut per haul. No statistics would need to be involved, and the monthly bycatch would be classified

as "dirty" with absolute certainty. In such case, no inferences need to be made about the unsampled hauls with respect to the incentive program.

It may be argued that these alternative definitions of "dirtiness" do not reward efficient vessels. That is, vessels that catch more groundfish per unit bycatch are not rewarded. There are merits to this argument. On the other hand, the fishery is shut down when the bycatch caps are reached, regardless of efficiency. So, it could be argued that bycatch of prohibited species reduces fishing time incrementally, regardless of vessel efficiency. However, if desired, efficient vessels that install scales on their vessels could be rewarded in that their "cleanliness" could be measured by classical bycatch rates (weight of bycatch per weight of all species in the catch samples expanded to the total haul) rather than bycatch levels.

Recommendations

1. Basket sampling of unsorted catches will likely result in substantial bias in estimates of bycatch rates for individual hauls. Also, bias in estimates of monthly bycatch rates may result. Further, expanded estimates of total bycatch from small subsamples are less accurate than total enumerations achieved by whole haul sampling for vessels on which whole haul sampling is feasible. For these reasons, it may be preferable to consider an alternative sampling plan rather than basket sampling. Because bycatch of rare species is to be estimated, it may be preferable to use a whole haul sampling plan when feasible or a modified whole haul sampling plan in all other cases.

2. Pope's (1963) advice is pertinent: *"Whichever method of sampling is used it is essential that it be subjected to a number of tests to ascertain its accuracy, for, only when unbiasedness has been proved should a method be properly adopted."* Data from the 1990 observer program should be analyzed. Bycatch estimates from vessels sampled by basket sampling and those sampled by whole haul methods could be compared for bias within time/area strata. Ideally, new experiments should be conducted on chartered commercial vessels in which basket samples and whole haul samples are collected simultaneously on each haul. The accuracy of the basket sampling plan should be fully evaluated for bias. Analyses should determine the sample size required to estimate bycatch rates with some fixed level of precision. An analysis should be conducted to determine the statistical power of tests of differences between monthly bycatch rates of individual vessels and the bycatch standard.

3. "Dirtiness" could be defined in some way other than bycatch rate in terms of weight of prohibited species per unit weight of all species. A rate could be calculated based on bycatch per haul or bycatch per unit time or some other easy-to-measure parameter. Alternatively, bycatch could be measured, not as a rate, but as a count of the number (or percentage) of a vessel's monthly hauls

in which bycatch of prohibited species exceeded some specified acceptable level. Other alternatives could be explored.

Acknowledgments

The author thanks Dave Carlile and Dave Ackley for constructive comments on this report, and numerous helpful discussions on commercial catch sampling problems.

Literature Cited

Cochran, W.G. 1977. Sampling techniques. John Wiley and Sons, New York.

May, A.W., and V.M. Hodder. 1966. Deck sampling of research vessel catches. Journal of the Fisheries Research Board of Canada 23:1083-1088.

NMFS (National Marine Fisheries Service). 1991. Groundfish of the Gulf of Alaska, Groundfish fishery of the Bering Sea and Aleutian Islands area: interim final rule; request for comments. Federal Register, draft notice.

Paloheimo, J.E., and L.M. Dickie. 1963. Sampling the catch of a research vessel. Journal of the Fisheries Research Board of Canada 20:13-25.

Pope, J.A. 1963. Sampling catches at sea. International Commission for Northwest Atlantic Fisheries, Special Publication 5:180-184.