A comparison of portside and at-sea sampling methods of estimating bycatch in the Atlantic herring fishery

Micah Dean
Herring PDT
May 2011

MarineFisheries
Commonwealth of Massachusetts

Introduction

There are three primary gear types used to target Atlantic herring in the United States: bottom trawls, midwater trawls and purse seines (Table 1). Estimates of bycatch in this fishery are primarily derived from data collected at-sea under the Northeast Fisheries Observer Program (NEFOP). On bottom trawl vessels, bycatch species are typically sorted from the catch on deck and are discarded at sea. NEFOP samplers often achieve a census of the bycatch before it is thrown overboard and as a result there is essentially no variance surrounding the estimate of bycatch on a trip. The other two gear types are considered “high-volume” fisheries and bycatch species usually remain mixed with the catch as it is pumped from the net to the hold, as it is impractical to sort through such large catches at sea. As such, sea samplers estimate bycatch by taking a systematic sample of the catch as it is pumped onboard. While this method of sampling provides a less precise estimate of bycatch than a census of discards, the fact that bycatch species are retained presents an opportunity to also sample the catch when it is offloaded at port.

Table 1. Average trip information for vessels participating* in the Atlantic herring fishery in 2010, as reported to the National Marine Fisheries Services (NMFS) on Vessel Trip Reports (VTRs).

Gear Type	Total number of trips	Median catch per trip (Kg all species)	Median tows per trip	Median trip length (days)
Purse seine	163	31,752	2	1
Midwater trawl	350	145,423	2	3
Bottom trawl	207	2,472	3	1

* *vessels with A,B,C or D herring permits not declared out of the fishery via VMS.*

In an effort to increase the number of trips sampled and thereby reduce the uncertainty surrounding fishery-wide estimates of bycatch, portside sampling programs have been initiated in Massachusetts and Maine. An obvious prerequisite to combining these portside and sea-sampling data is the comparability of the sampling programs. An initial comparison found relatively poor agreement between the two methods and raised concerns over the ability of either program to estimate bycatch in this fishery (Appendix A). The objective of this report is to explain the source of this previous disagreement, and to provide an updated comparison of the different programs. At-sea and portside sampling protocols were compared using a simulation model as well as with empirical data from herring trips sampled under both programs. This comparison focuses on midwater trawl vessels, as they present the greater challenge in sampling at-sea would therefore benefit the most from additional portside sampled trips.

Methods

Simulation

The contents of a typical, yet simplified hold of a midwater trawl herring vessel were simulated in the R software package by assembling an array of individual fish caught from three tows, totaling 150 mt in weight. The tows were of equal size (50 mt), but to evaluate the sensitivity of each sampling protocol to non-randomly distributed bycatch, two different scenarios were evaluated: 1) similar bycatch per tow and 2) dissimilar bycatch per tow. To simplify the comparison between protocols, the simulated hold contained only 3 species: Atlantic herring (target species), river herring[1] (higher abundance bycatch species) and whiting (lower abundance bycatch species). The hold under each scenario contained a similar amount of each species and differed only in the concentration of bycatch species in each tow (Table 2).

Table 2. Percent of target and bycatch species by weight in each tow under each simulation scenario.

Similar Tows Scenario

	Atlantic herring	river herring	whiting
Tow 1	98.9%	1.0%	0.1%
Tow 2	98.9%	1.0%	0.1%
Tow 3	98.9%	1.0%	0.1%
Total	98.9%	1.0%	0.1%

Dissimilar Tows Scenario

	Atlantic herring	river herring	whiting
Tow 1	99.89%	0.10%	0.01%
Tow 2	97.40%	2.40%	0.24%
Tow 3	99.41%	0.50%	0.05%
Total	98.9%	1.0%	0.1%

The number of individuals of each species (n_s) in a given tow was determined by multiplying the percent contribution of that species (p_s) by the weight of the tow (w_{tow}), and dividing by the average individual size ($\overline{w}_s$ - Figure 1).

[1] For the purposes of estimating bycatch in the Atlantic herring fishery, alewife (*Alosa pseudoharengus*) and blueback herring (*Alosa aestivalis*) are typically grouped together and referred to as "river herring."

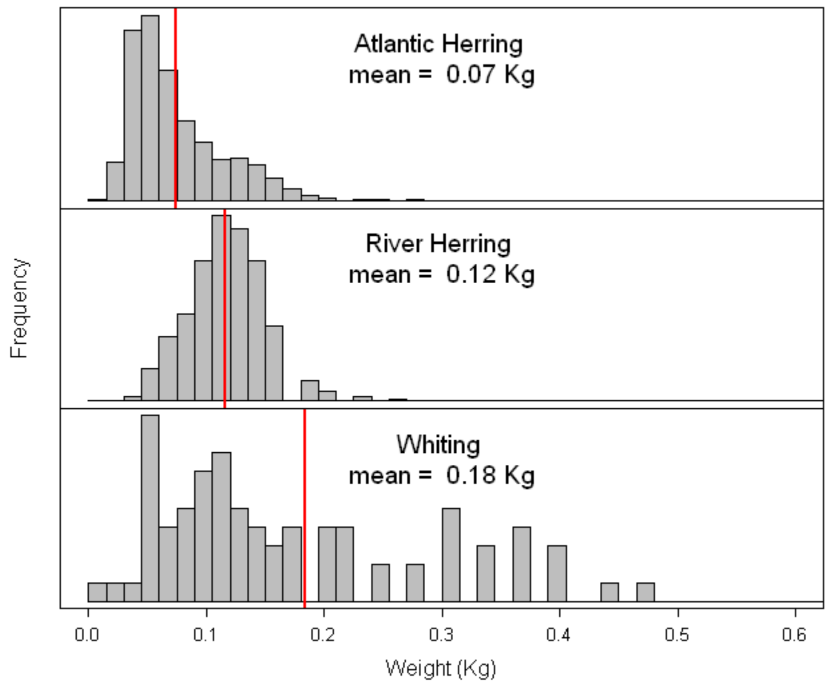

Figure 1. Frequency distributions of individual fish weights from observed herring trips in 2010.

A tow was represented by creating an array of individual fish of length n_{tow} ($\sum n_s$). For each bycatch species, a random sample of n_s individuals was selected from this array and designated as species *s*. The remaining individuals were designated as the target species, Atlantic herring. Individual fish were assigned random weights by sampling with replacement from the distribution of weights observed at-sea in 2010 for that species (Figure 1). The three tow arrays were joined into a single array to represent the hold. The bycatch distribution from each tow was kept intact when assembled into the hold array (i.e. no mixing) to mimic the process of pumping each tow's catch into a separate tank onboard the vessel. The total weight of each bycatch species was estimated by sampling from the hold array using four different protocols: 1) at-sea sampling: (AS) 2) portside unsorted sampling: (PU) 3) portside sorted sampling: (PS) and 4) portside lot sampling: (PL).

At-Sea Sampling (AS)

In high-volume fisheries, at-sea samplers typically take a systematic sample of 10 "baskets" from each tow to describe the species composition of the catch. The contents of each ~30 kg basket are sorted by species and weighed. The proportion of each species from the basket sample is then multiplied by the captain's estimate of tow size to achieve the amount of each species per tow. The total amount of bycatch for the trip is the sum of the bycatch estimates from each tow (Table 3). This sampling process was simulated by taking a systematic sample of 10 individuals from each tow array, with a random starting point. Each selected individual represented the beginning of a basket, and every fish following the initial one was added to the basket until the weight of the basket exceeded 30 kg.

Portside Unsorted Sampling (PU)

During the offload process, vessels typically pump their catch through a “de-watering box” as it is transferred from the hold to trucks or vats onshore. Portside samplers take a systematic basket sample of the catch as it is pumped through the de-watering box. In this case, the sampling interval is determined by the amount of time required to sort and weigh the contents of a basket, which on average is about seven minutes. Since the amount of time required to offload a vessel can be 24 or more hours, samplers take occasional breaks from sampling and are typically working for approximately 75% of the pumpout. The total amount of each bycatch species from the basket sample is expanded to the entire hold using the captain’s estimate of the trip size (Table 3). This sampling protocol was simulated in a manner similar to the method of creating baskets under the AS protocol; However, in this case a single systematic sample of 32 baskets was taken from the hold array, based on an estimated 5 hour pumpout for a trip of this size (5 hours x 75% sampling time / 7 min per basket = 32 baskets).

Portside Sorted Sampling (PS)

The majority of Atlantic herring landings are sold as bait for the American lobster fishery, and as such are typically transferred directly from vessels to waiting trucks or vats. Alternatively, some herring are packaged, frozen and sold as food, often in international markets. Herring destined to be food are pumped from the vessel onto a conveyor belt at a shoreside facility where as many as 8 to 12 personnel (i.e. “pickers”) separate bycatch species from the catch prior to packaging. The bycatch, as well as any damaged Atlantic herring, are sent to a central vat via a series of chutes. Portside samplers systematically sample the flow of bycatch to this vat via baskets in a manner similar to unsorted sampling. However, since the bycatch is more concentrated at this sampling location, processing the basket contents is more laborious and the average sampling interval is approximately 15 minutes. Likewise, samplers often take longer breaks and are working for approximately 50% of the pumpout. The total amount of each bycatch species from the basket sample is expanded to the entire hold using the total amount of bycatch (and damaged Atlantic herring) sorted to the vat (Table 3). To simulate this protocol, 5% of the target species in the hold array were randomly designated as damaged. The bycatch vat was represented by extracting a subset of the hold array that contained all bycatch species and damaged Atlantic herring, maintaining the original order from the hold array. A single 10-basket systematic sample from this vat array was taken, using the method described under the AS protocol (5 hours x 50% sampling time / 15 min per basket = 10 baskets).

Portside Lot Sampling (PL)

Herring that are sold as lobster bait are often pumped directly into trucks and driven over land to dealers in Maine. When it is not possible for portside samplers to intercept and sample the vessel during the pumpout, the contents of bait trucks are often sampled as they are unloaded at the retail dealer in Maine. In this case, a systematic sample of baskets is taken from the contents of the truck as it is emptied. Typically, 1-3 trucks from a vessel are sampled and together are referred to as a "lot." On average, a 16-basket sample is taken from each truck. The total amount of bycatch from a trip is calculated by expanding the basket sample first to the entire lot, and then from the lot to the entire hold (Table 3). At times, conditions permit sorting all bycatch species from the lot as it is offloaded. In these cases, a census of bycatch from the lot is achieved. However, to simplify the comparison with other protocols, the systematic basket-sampling approach was used to represent PL sampling in the simulation. The hold array was broken up into eight sections to represent individual trucks, two of which were randomly selected for sampling. A 16-basket systematic sample was made from each selected truck, using the method described under the AS protocol.

A total of 1,000 iterations of each sampling protocol were made from the simulated hold under each scenario. The distribution of bycatch estimates from all sampling iterations was used to describe the accuracy and precision of each protocol. The mean estimate and coefficient of variation (CV) for each bycatch species were compared across the four protocols and two scenarios

A second simulation experiment was conducted to illustrate the effect of sample size and bycatch rarity on the precision of bycatch estimates as well as the ability to detect a bycatch species. In this case, a hold was simulated that contained four randomly-distributed bycatch species at various densities (1%, 0.1%, 0.05%, and 0.01%). A single systematic sample with random starting point was taken from this hold at various sample sizes (10, 25, 50 and 100 baskets). This sampling routine was iterated 1000 times, and the CV of the estimates, as well as the percent of estimates that were zero were compared.

Table 3.

Sampling protocol	Method used to estimate bycatch and variance	
At-Sea Sampling (AS)	$r_{s,tow} = \frac{\sum w_{s,i}}{\sum w_i}$ $w_{s,tow} = r_{s,tow} w_{tow}$ $w_s = \sum w_{s,tow}$ $V(w_{s,tow}) = N^2\left(\frac{N-n}{N}\right)\left(\frac{s^2}{n}\right)$ $s^2 = \frac{\sum (w_{s,i} - r_{s,tow} w_i)^2}{n-1}$ $V(w_s) = \sum V(w_{s,tow})$	Where, $r_{s,tow}$ = ratio of species *s* in tow $w_{s,i}$ = weight of species *s* from basket *i* $w_{s,tow}$ = weight of species *s* from tow w_s = weight of species *s* in the hold w_i = weight of basket *i* N = number of possible baskets from tow n = number baskets sampled from tow
Portside Unsorted (PU)	$w_s = \frac{N}{n}\sum w_{s,i}$ $V(w_s) = N^2\left(\frac{N-n}{N}\right)\left(\frac{s^2}{n}\right)$	Where, w_s = weight of species *s* in the hold $w_{s,i}$ = weight of species *s* from basket *i* N = number of possible baskets from hold n = number of baskets sampled from hold s^2 = sample variance of species *s* from baskets
Portside Sorted (PS)	$w_s = \frac{N}{n}\sum w_{s,i}$ $V(w_s) = N^2\left(\frac{N-n}{N}\right)\left(\frac{s^2}{n}\right)$	Where, w_s = weight of species *s* in the hold $w_{s,i}$ = weight of species *s* from basket *i* N = number of possible baskets from bycatch vat n = number baskets sampled from bycatch vat s^2 = sample variance of species *s* from baskets
Portside Lot (PL)	$w_{s,lot} = \frac{N}{n}\sum w_{s,i}$ $w_s = \left(\frac{M}{m}\right) w_{s,lot}$ $V(w_{s,lot}) = N^2\left(\frac{N-n}{N}\right)\left(\frac{s^2}{n}\right)$ $V(w_s) = \left(\frac{M}{m}\right)^2 V(w_{s,lot})$	Where, w_s = weight of species *s* in the hold $w_{s,i}$ = weight of species *s* from basket *i* N = number of possible baskets in lot n = number of baskets sampled from lot M = number of possible trucks from hold m = number of trucks sampled from hold s^2 = sample variance of species *s* from baskets

Empirical Data

A total of 30 midwater-trawl herring trips from 2010-2011 were identified as being sampled by both at-sea and portside methods. Twenty-four trips were sampled by PU methods and six trips were sampled by PS methods. Five trips were sampled by more than one portside method (PU and PL). Estimates of bycatch for six common species (river herring, whiting, American shad, butterfish, haddock and spiny dogfish) were calculated for each trip and compared across sampling methods. The variability of the bycatch observed in the basket sample data was used to estimate the variance surrounding the bycatch estimate for each trip (Table 3). The amount of agreement between at-sea sampling and port sampling was evaluated in two different ways: 1) the ability to detect bycatch species (i.e. presence-absence) and 2) significant differences in bycatch estimates.

There are four possible outcomes when comparing two sampling protocols' ability to detect a bycatch species: 1) present in both (++); 2) present in neither (--); 3) present in portside sampling, absent from sea sampling (P+); and 4) present in sea sampling, absent from portside sampling (S+). For each of the six species, a matching coefficient was calculated by dividing the number of trips that fell into the first two categories (++,--) by the total number of trips sampled. Because some species are infrequently encountered by this fishery, a high proportion of "double negative" (--) trips could yield an unrealistically high matching coefficient. To account for this, a second matching coefficient was calculated that omits the "--" trips from both the numerator and denominator.

A modified two sample *t*-test assuming unequal variances (i.e. Welch's test) was used to test for a significant difference ($\alpha = 0.05$) between portside and at-sea sampling estimates of bycatch. This test typically relies on the sample means ($\bar{x}$) and sample variances (s^2) to calculate the t statistic:

Eq. 2)
$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

However, since we are comparing a variety of protocols with different *types* of samples (e.g. sorted vs unsorted), the sample means and variances were replaced with the total estimated bycatch per trip (w), and the variance of those estimates ($V(w)$):

Eq. 3)
$$t = \frac{w_1 - w_2}{\sqrt{V(w_1) + V(w_2)}}$$

To evaluate how often bycatch is non-randomly distributed in the catch, a one-sided "runs" test was performed on the series of basket observations for a single species (river herring) for each sampled trip. This test relies on the calculation of an expected number of "runs" given the number of observations at two levels, a "run" being a sequence of adjacent observations at the same level. In this case, the two levels were defined as: 1) above the mean value and 2) below the mean value. If the number of observed runs for a trip was significantly lower than the expected value ($\alpha = 0.05$), it was considered to have non-randomly distributed bycatch.

Results

Simulation

AS, PU and PL sampling achieved comparable levels of precision under the similar tows scenario, with a CV of approximately 0.1 for river herring and 0.6 for whiting (Figure 2). PS sampling was identified as being the most precise, with a CV roughly half of the other protocols. AS, PU and PS sampling performed equally well under the dissimilar tows scenario, indicating these protocols are robust to non-random distributions of bycatch (Figure 3). On the other hand, PL sampling performed very poorly under the dissimilar tows scenario, yielding a CV of 0.59 for river herring (a 500% increase) and a CV of 0.91 for whiting (an 80% increase). Additionally, the PL protocol failed to detect any of the less abundant whiting 9% of the time under the dissimilar tows scenario. None of the protocols were found to be biased, achieving mean estimates within 3% of the true value under either scenario.

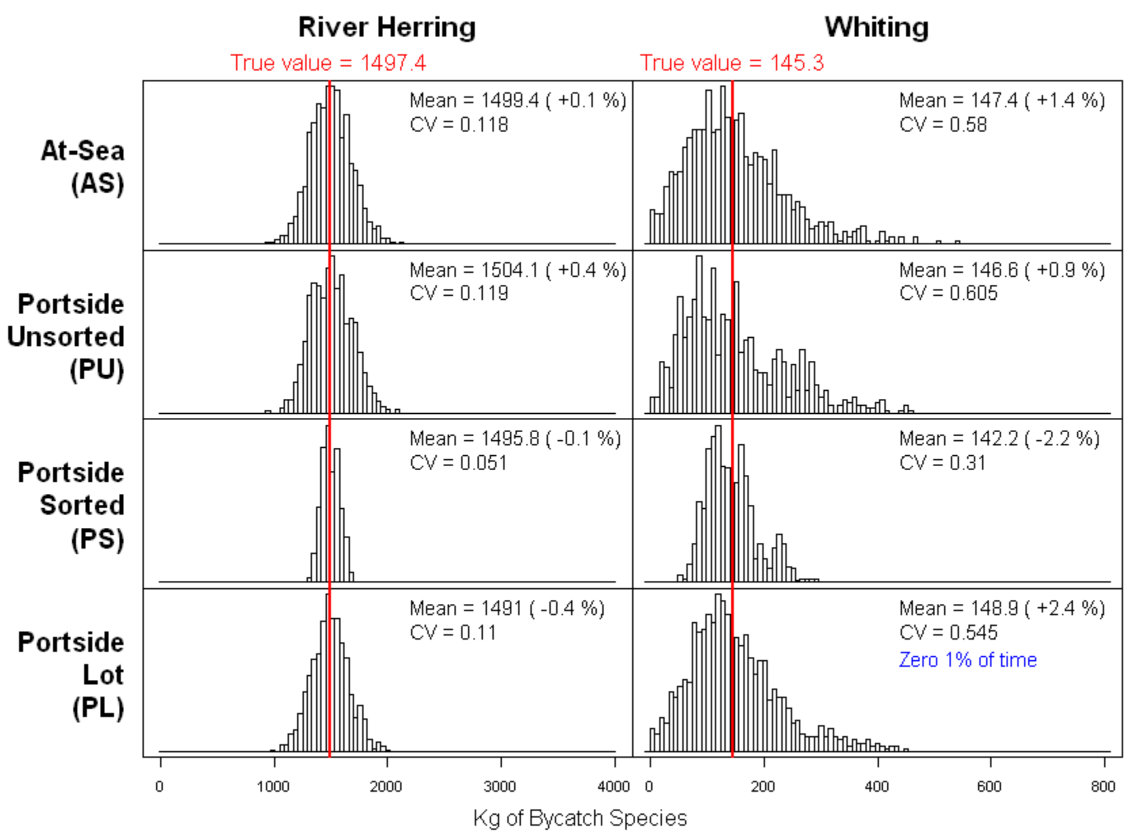

Figure 2. Distribution of bycatch estimates from 1000 iterations of sampling the simulated hold under the similar tows scenario.

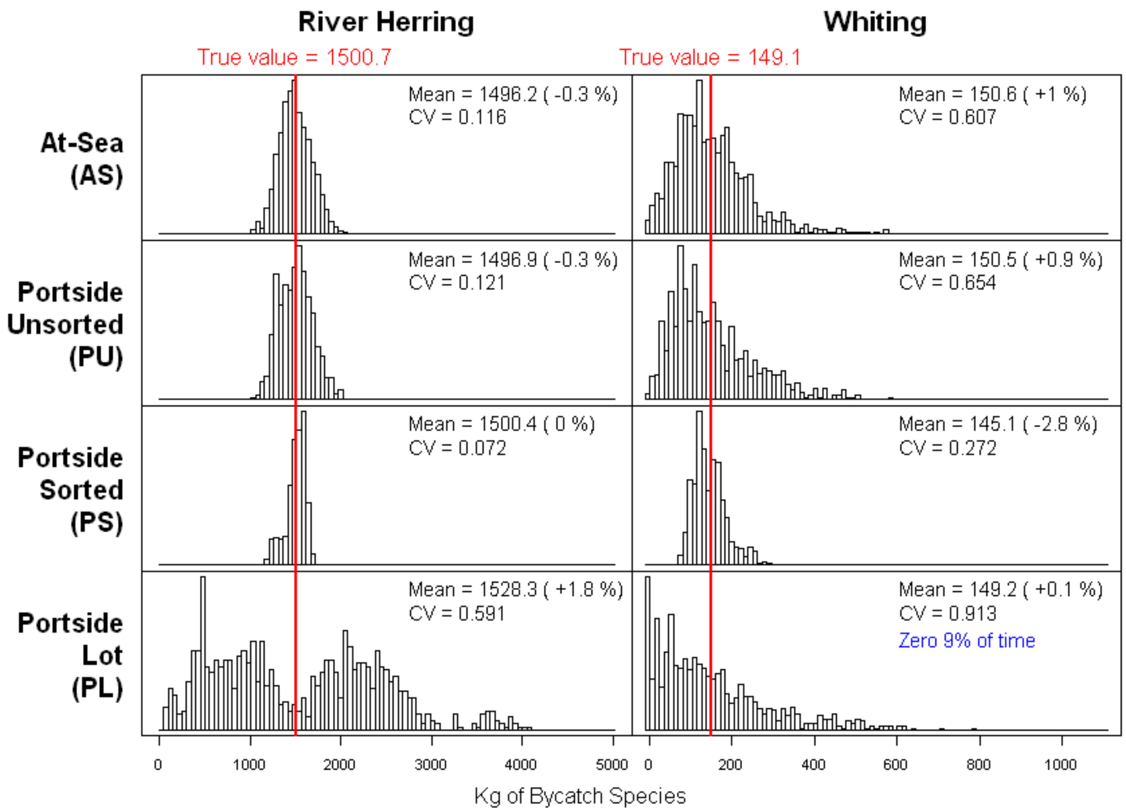

Figure 3. Distribution of bycatch estimates from 1000 iterations of sampling the simulated hold under the dissimilar tows scenario.

For the second simulation experiment, results indicated that both sample size and bycatch rarity have a strong influence on the precision of the estimate and the ability to detect bycatch species (Figure 4). For all levels of bycatch rarity, the CV of the estimate increased by a factor of 3 when the sample size was dropped from 100 baskets to 10 baskets. Likewise, the CV of the estimate increased by a factor of 10 when the rarity of the bycatch dropped from 1% of the catch to 0.01%. The ability to detect the rarest bycatch was very low at the smallest sample size, with 71% of the sampling iterations failing to detect a single bycatch individual.

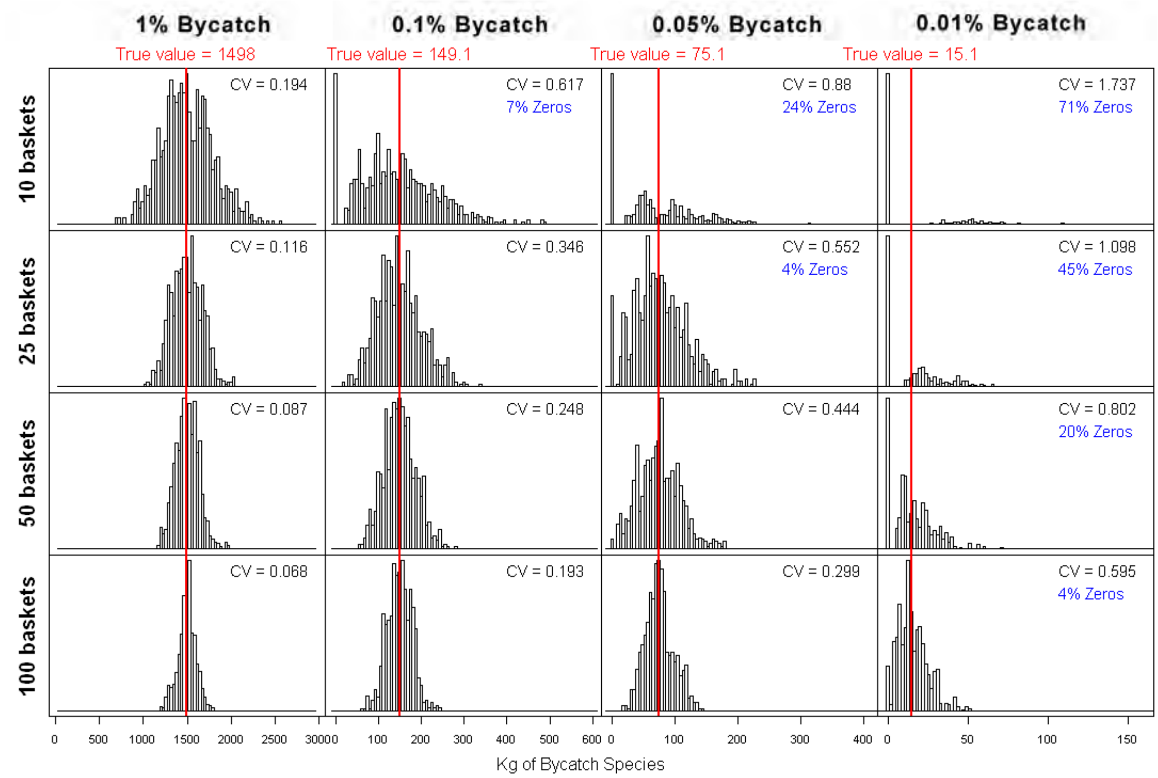

Figure 4. Distribution of bycatch estimates at four levels of bycatch rarity from 1000 iterations of a single systematic sample of various sizes.

Empirical Data

PU sampling achieved a 75% to 92% agreement with AS in the ability to detect the six bycatch species (Table 4). When the “double negative” (--) tows were omitted, the amount of agreement dropped substantially for some species (e.g. spiny dogfish, butterfish), yet remained high for others (e.g. river herring, silver hake). The *amount* of bycatch estimated also agreed well between PU and AS sampling. For the six species considered, significant differences in bycatch amount were detected on only 8 to 25% of the trips (75 to 92% agreement). All of these significant differences resulted from one protocol “missing” the bycatch species (i.e. P+ or S+).

Table 4. Number of trips that detected bycatch species under PU and AS and the percent agreement between the two protocols (top). Number of trips with a significant difference (alpha = 0.05) in bycatch amount of six species, between AS and PU methods (bottom).

Presence-Absence

	Agree		Disagree		% Agree	% Agree (omit --)
	++	- -	P+	S+		
River Herring	18	3	1	2	88%	86%
Butterfish	5	15	3	1	83%	56%
Silver Hake	12	9	1	2	88%	80%
Spiny Dogfish	2	19	3	0	88%	40%
Haddock	6	16	2	0	92%	75%
American Shad	7	11	4	2	75%	54%

Significant Differences

	No Sig Diff		Sig Diff			% No Sig Diff	% No Sig Diff (omit --)
	++	- -	++	P+	S+		
River Herring	18	3	0	1	2	88%	86%
Butterfish	5	15	0	3	1	83%	56%
Silver Hake	12	9	0	1	2	88%	80%
Spiny Dogfish	2	19	0	3	0	88%	40%
Haddock	6	16	0	2	0	92%	75%
American Shad	7	11	0	4	2	75%	54%

PS sampling achieved 50% to 83% agreement with AS in the detection of bycatch species (Table 5). However, for some species the agreement dropped to 0% when "--" trips were omitted (i.e. spiny dogfish, haddock). In terms of the quantity of bycatch estimated, PS sampling had limited agreement with AS sampling: 33%-83% of the trips sampled had significant differences between the protocols (17%-67% agreement). However, much of that agreement came from "double negative" trips and once they were removed, the amount of agreement dropped to 0%-50%.

Table 5. Number of trips that detected bycatch species under PS and AS and the percent agreement between the two protocols (top). Number of trips with a significant difference (alpha = 0.05) in bycatch amount of six species, between AS and PS methods (bottom).

Presence-Absence

	Agree		Disagree		% Agree	% Agree (omit --)
	++	- -	P+	S+		
River Herring	5	0	1	0	83%	83%
Butterfish	2	2	2	0	67%	50%
Silver Hake	4	0	2	0	67%	67%
Spiny Dogfish	0	3	3	0	50%	0%
Haddock	0	4	2	0	67%	0%
American Shad	3	0	3	0	50%	50%

Significant Differences

	No Sig Diff		Sig Diff			% No Sig Diff	% No Sig Diff (omit --)
	++	- -	++	P+	S+		
River Herring	2	0	3	1	0	33%	33%
Butterfish	1	2	1	2	0	50%	25%
Silver Hake	1	0	3	2	0	17%	17%
Spiny Dogfish	0	3	0	3	0	50%	0%
Haddock	0	4	0	2	0	67%	0%
American Shad	3	0	0	3	0	50%	50%

The comparison of PL to AS sampling was restricted to only three species, as none of the other common bycatch species were encountered under either protocol. PL sampling achieved 20% to 80% agreement with AS in the detection of those three bycatch species (Table 6). Similarly, significant differences between the two protocols were found on 20% to 80% of the trips. While it is difficult to draw conclusions from so few trips sampled, it appears that the amount of agreement between PL and AS sampling was highly variable.

Table 6. Number of trips that detected bycatch species under PL and AS and the percent agreement between the two protocols (top). Number of trips with a significant difference (alpha = 0.05) in bycatch amount of six species, between AS and PS methods (bottom).

Presence-Absence

	Agree		Disagree		% Agree	% Agree (omit --)
	++	- -	P+	S+		
River Herring	3	1	0	1	80%	75%
American Shad	0	1	1	3	20%	0%
Spiny Dogfish	1	3	1	0	80%	50%

Significant Differences

	No Sig Diff		Sig Diff			% No Sig Diff	% No Sig Diff (omit --)
	++	- -	++	P+	S+		
River Herring	2	1	1	0	1	60%	50%
American Shad	0	1	0	1	3	20%	0%
Spiny Dogfish	1	3	0	1	0	80%	50%

Of the 30 trips that were sampled both at-sea and portside, 11 trips (37%) were found to have a significantly non-random distribution of river herring in the catch when sampled at-sea (Figure 5). Similarly, nine trips (30%) were found to have non-randomly distributed river herring when sampled portside. Six of these portside trips were also found to be non-random at-sea, indicating that the distribution of bycatch at-sea often determines the distribution of bycatch seen during the offload at port.

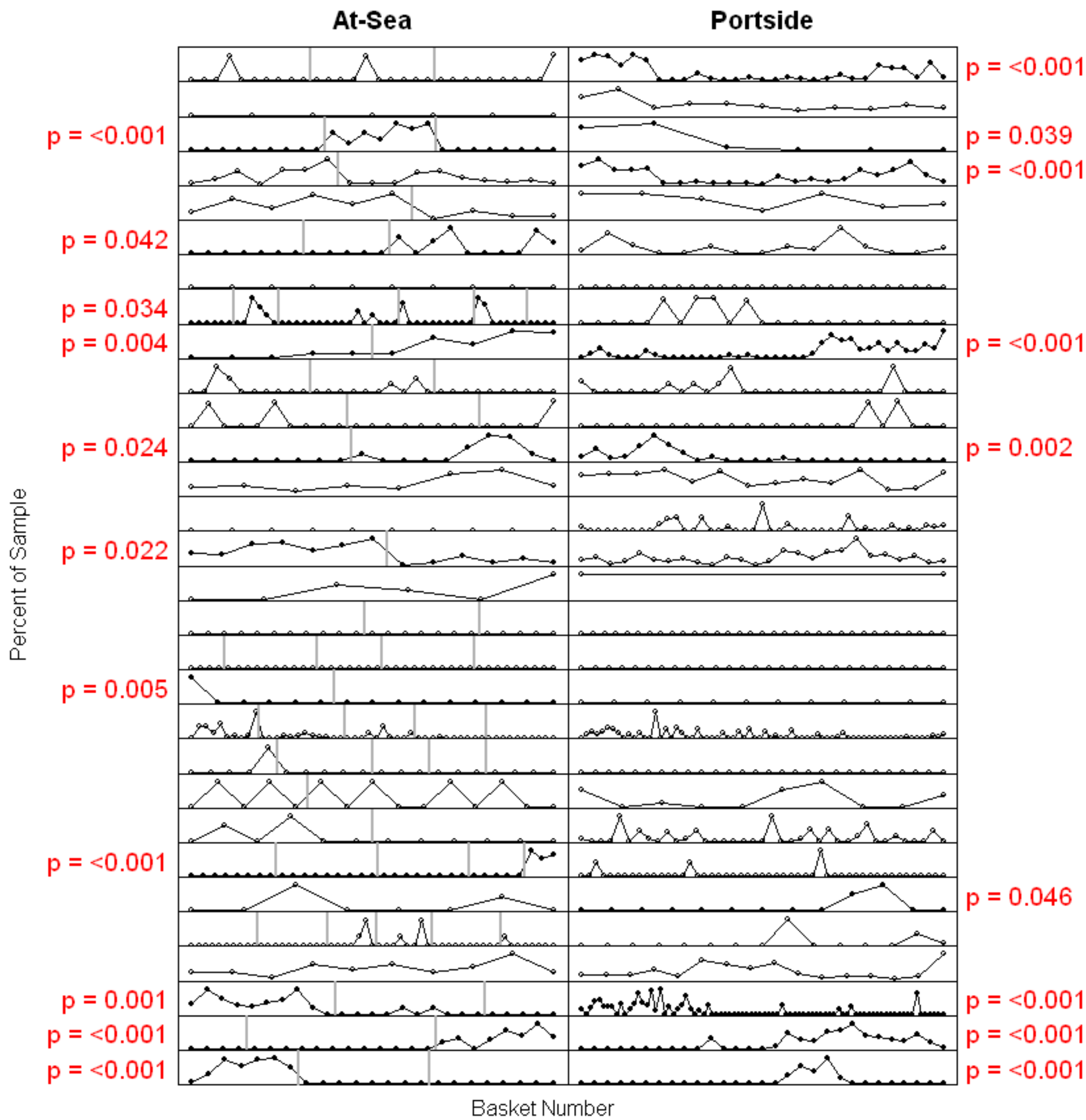

Figure 5. Sequence of river herring observations from the basket samples taken at-sea (left) and portside (right) for each co-sampled trip. The vertical gray lines on the left pane indicate the separation between tows. Trips that were found to have a non-random distribution of river herring are identified by the *p*-value of the runs test in the margin.

Discussion

The results of the simulation experiment indicate that AS and PU sampling should estimate the amount of bycatch on a herring trip with equal precision. However, the precision surrounding these bycatch estimates is primarily determined by the sample size (i.e. the number of baskets), and each protocol arrives at this number differently. The configuration of the hold in the simulation experiment (3 tows with 150 mt total landings) was intentionally designed to create roughly equivalent sample sizes under each protocol. In reality, most herring trips would likely end up with different sample sizes if sampled at-sea or portside. On trips with relatively low landings from a high number of tows, AS sampling will achieve the larger sample size, since AS protocol requires a fixed number of baskets per tow. On the other hand, trips with high landings from few tows will yield a higher sample size under PU sampling, as number of samples taken is a function of pumpout time under the PU protocol. For the 24 trips that were sampled by both methods, 75% had a higher sample size under PU sampling. As a result, the average CV from PU sampling was 42% less than that achieved under AS on the same trips. Despite these dissimilar sample sizes, the empirical data appear to corroborate the results of the simulation, showing little disagreement between PU and AS sampling. None of the trips where both protocols detected a species had significant differences in bycatch estimates. Most of the trips where PU and AS did disagree fell into the "P+" category, or where bycatch was observed in PU but not observed in AS. This is likely a symptom of the larger sample sizes under portside sampling. From the second simulation experiment, it is clear that the likelihood of not detecting a rare bycatch species is far greater at lower sample sizes.

The simulation identified PS sampling as having the highest precision, while being just as accurate as PU and AS methods. Unfortunately, this did not appear to be the case when the empirical data were examined. Although only six trips were sampled by both PS and AS methods, it was clear that the two methods had very poor agreement. Four trips (67%) showed a significant difference in the estimated amount of river herring bycatch. This high amount of disagreement was a surprising find and indicates the presence of a strong bias in PS sampling. It is unlikely that AS protocol is biased, since it is corroborated by the good agreement with PU sampling. A potential source of bias from PS sampling could be that pickers are "missing" bycatch species as they pass by on the conveyor belt. The PS method of sampling assumes that all bycatch are separated from catch and sent to a vat to be sampled from. If a large portion of a particular bycatch species is missed on the picking line, the PS method will significantly underestimate the amount of bycatch. In fact, three of the four trips that had a significant difference in river herring bycatch had a lower estimate under PS sampling. Occasionally, a few random boxes of packaged Atlantic herring are opened for quality control purposes and examined for missed bycatch species. However, since bycatch species such as river herring often account for less than 1% of the weight of trip, the chance of encountering the occasional 'missed' river herring in a randomly selected box is very small. Likewise, if a fraction of 1% of the packaged product is of a different herring species, it unlikely that it would lead to customer complaints.

Although apparently biased, the average precision surrounding the PS estimates of bycatch was 40% less than that of AS sampling on the same trips. Also, PS sampling was far better at identifying the presence of rare bycatch species than AS, with 13 instances of a species being identified under PS sampling but not under AS (P+). There were no cases of a species being identified under AS but not under PS (S+). If it is possible to identify and correct the source of the sampling bias, the PS protocol has the potential to provide the most precise estimate of bycatch for this fishery.

The simulation experiment identified PL sampling as being roughly equivalent to PU and AS sampling for trips that had randomly distributed bycatch in the hold (i.e. similar tows scenario). However, PL sampling was particularly vulnerable to non-random bycatch distribution. Intuitively, this vulnerability is caused by focusing all of the sampling on a small portion of the hold. If bycatch are more concentrated in the trucks selected for sampling, the resulting bycatch estimate will be too high. Conversely, if trucks with less bycatch are selected, the bycatch estimate will be too low, or none will be detected at all. The other protocols distribute the sampling effort across the entire catch, and are therefore more robust to non-randomly distributed bycatch. From the results of the runs test, it appears that non-random bycatch is fairly common in this fishery, with more than a third of the examined trips identified as having river herring non-randomly distributed in the catch. A previous comparison of portside and at-sea bycatch estimates relied heavily on data from PL sampling (Appendix A) and yielded relatively poor agreement between portside and at-sea estimates of bycatch. It appears the frequency of non-random bycatch coupled with the vulnerability of PL sampling to this phenomenon is the primary cause of the disagreement found in that investigation. It is important to note that sampling the offload from trucks is not inherently flawed, and could provide a reliable estimate of bycatch if it is possible to sample *all* the trucks from a particular trip. In many cases this may not possible, as trucks filled from a single hold are often destined for multiple locations. In any case, portside resources should be directed to ensure that the entire catch from a trip is available for sampling.

Currently, estimates of bycatch in this fishery are derived from AS sampling alone. The high amount of agreement between PU and AS sampling found in this investigation lends credibility to both programs and it seems reasonable to combine bycatch estimates from trips sampled under either protocol. Incorporating PU sampled trips could reduce the CV on the fishery-wide estimate of bycatch by dramatically increasing the sample size, particularly for areas and gears with limited AS sampling coverage. In addition, PU sampling can be a far more efficient use of limited resources than AS sampling: Consider the midwater-trawl trip from the simulation experiment, with 150 mt of landings that takes approximately 5 hours to offload. Depending on the fishing area and port of landing, this might be a four to six day trip. The total cost of sampling that trip portside would be approximately \$350 (2 samplers at ~\$35 / hr for 5 hrs), whereas the AS cost would likely be \$5000 to \$7000 (1 sampler at ~\$1200 / day). If additional sampling resources are to be directed at this fishery beyond those already required by the NEFOP allocation algorithm (SBRM), portside unsorted sampling should be considered as it is a reliable cost-effective method of estimating retained bycatch.