SCIENTIFIC COMMITTEE

EIGHTEENTH REGULAR SESSION

Online meeting

10-18 August 2022

DESIGNING EM REVIEWING RATES FOR WCPFC FISHERIES

WCPFC-SC18-2022/ST-IP-06

SPC-OFP

Executive Summary

Electronic Monitoring/Observation of tuna fisheries in the western and Central Pacific Ocean is currently undergoing various trials throughout the region to ascertain how it can best complement existing at-sea observer programs. The types of electronic monitoring being considered typically includes both camera and metered observation systems (e.g. location and gear usage sensors). Camera systems typically include some review of the footage collected by an analyst once the vessel has returned to port. While reviewing the footage the analyst collects the relevant data in a manner equivalent to being on board the vessel at the time of fishing.

As footage can be viewed at speeds faster than real-time there are potential cost savings associated with less time required to view the entire footage of a fishing trip. Noting that an intent of electronic monitoring is to potentially ensure all vessels and all trips have video footage than can be monitored, a key question is what proportion of trips by each vessel and what proportion of a trip duration needs to be viewed in order to obtain the desired precision in the fisheries data fields for scientific purposes. Only viewing the required amount of footage by the analyst to achieve the desired precision present opportunities for considerable cost savings for implementing electronic monitoring programs.

SPC-OFP in collaboration with Patrick Cordue (Innovative Solution Limited) initiated analyses and simulations to estimate the precision for each observer data field associated with differing region, fleet and trip coverage rates for WCPFC longline and purse-seine fisheries. A copy of this report is attached to this Information Paper.

In this study, a general sampling scheme involving the random sampling of vessels, trips, and sets was investigated. The primary data were the observer and logbook data for the years 2016-2019 inclusive. Analytical equations were developed for the variance and coefficient of variation (CV) of an unbiased estimator of total catch for a given species, area, and timeframe.

It was found, when estimating total catch for a given species, that the simple approach of sampling a proportion of sets from every trip was almost always superior to any alternative approach. Even complex stratifications using information that would not generally be available at the time of stratification did not outperform the simple approach. This is because between-vessel and between-trip variation in catch can be large. When all trips are selected the between-trip and between-vessel variation in catch does not affect the variance of the estimator.

The main focus of the study was catch rather than nominal CPUE as the precision of estimators sampling sets was almost the same for both. This is because the CV for total hooks is small compared to the CV for total catch (when all trips are sampled). Intuitively this is clear as for a given trip there is unlikely to be much variation in the hooks deployed per set but the variation in catch per set (for a given species) can be very high.

The preliminary analyses in this report and the development of operating models to scale observer data up to the level of logbooks laid the foundation for the study (though ultimately the operating models were not important). From the preliminary analyses of longline fisheries, it was clear that for target species over a large area and an annual timeframe that sampling 10% of the sets on each trip would provide good precision for total catch (e.g., CV less than 10%). This is because of the large number of sets that occur each year in broadscale areas and because target species are caught on a majority of the sets. Indeed, for some target species, a sampling proportion of 5% was adequate.

For bycatch species, an estimator's CV is not always a good indicator of appropriate precision. For example, when there is a catch of 20 animals for a species it is not necessary to have a CV of 10% or less (e.g., a 95% CI of about 16-24). It is generally of little consequence whether there were 10, 20, or even 80 animals caught. Depending on the tolerance for precision, for infrequently caught species (5-20% of sets) a sampling proportion of 10-20% will generally be adequate. For the range of species looked at for longline, over shorter timeframes and smaller areas, a sampling proportion of 20% was also generally adequate (e.g., to estimate catch at a regional level). For purse seine catches a sampling proportion of 20% was also generally adequate.

For rarely caught species (less than 5% of the sets) the main consideration may be whether any of the species were caught or not. In this case, it was more important to consider the number of sets involved and the proportion of sets on which the species is generally caught (i.e. defining the threshold for the probability of incorrectly reporting a zero catch). For very rarely caught species it may be that to avoid falsely reporting a zero catch that most of the sets need to be sampled. Hopefully, in these cases the number of sets that need to be sampled can be reduced by restricting the area to the known habitat of the species.

It is recommended that SC18 note the framework used to simulate the "footage review rates" of potential electronic monitoring in the WCPFC fisheries for scientific data collection and the outcomes of the simulations undertaken:

- To maximise precision with minimize review rates for estimating total catch for a given species, the capacity to sample a proportion of sets from every trip was almost always superior to the alternative of a stratified sample of trips.
- Assuming a sufficient proportion of the fleet has EM, randomly selecting at least 10% of sets from each trip for analysis would be sufficient for the estimation of target species catch (with CVs <10%) at the WCPFC spatial scale.
- That increased sampling proportions will likely be required when good precision is required at a sub-regional level (as a rule of thumb 20% of sets sampled at random from each trip may be more appropriate).
- By design, a higher proportion of sets may need to be sampled to meet specific criteria (e.g., to determine if a species of special interest has been caught or not).

Sampling of electronic monitoring data for longline and purse seine fisheries

P.L. Cordue

13 December 2021

ISL Client Report for SPC

Introduction

It is anticipated that electronic monitoring of the tuna fleet in WCPFC area will become widespread in the next few years. This study looks at the precision of fisheries statistics associated with different sampling strategies for EM analysts.

The current recommendation for EM analysts, for longline data, is to view a randomly selected proportion of sets for each trip and to target 20% of sets for non-domestic vessels with a minimum of 10% of sets viewed for domestic vessels (FFA 2020). This recommendation is based on the work of Lawson (2003, 2004) and Peatman & Nicol (2020). Lawson noted that the CV of an estimator of nominal CPUE decreased slowly as a function of the proportion of sets sampled from about 20%. Peatman & Nicol (2020) showed that sampling a proportion of sets from each trip was superior to sampling all sets from a proportion of trips.

In this report, a general sampling scheme involving the random sampling of vessels, trips, and sets is investigated. The focus of the work is on the longline data, but purse seine trips are also analysed.

Methods

Preliminary analyses

The primary data studied were the observer and logbook data for the years 2016-2019 inclusive with the exception of some domestic fisheries which are usually excluded from area-wide analyses. Observer and logbook data were used for longline fisheries and for purse seine fisheries observer data were used (as there is very high coverage). The initial and main focus of the study were the longline data. Purse seine data were considered later in the study.

The data were loaded into R and functions to link, analyse, and plot the data were developed. Three alternative sampling schemes were considered for preliminary analyses:

Sampling sets: a proportion of the sets on every trip by every vessel are sampled randomly within trip

Sampling trips: a proportion of the trips **across** all vessels are sampled randomly and every set on a selected trip is sampled

Sampling trips **within** vessel: a proportion of the trips by each vessel are sampled randomly and every set on a selected trip is sampled.

The statistics considered in the preliminary analyses were total catch numbers (for a given species) and total number of hooks (hooks observed for the observer data and hooks set for the logbook data). Species and latitudes considered were albacore tuna (ALB) (30° S to 10° S), bigeye tuna (BET), skipjack tuna (SKJ), yellowfin tuna (YFT), blue marlin (BUM), and olive Ridley turtle (LKV) (10° S to 20° N)

For each sampling scheme the proportion sampled, p, was varied from 0.01 to 1 in steps of 0.01. The CV of the estimator of the total (catch or hooks) was calculated analytically for each sampling scheme (see Appendix 1). The expected number of sets sampled was also calculated analytically for each value of p and each sampling scheme (see Appendix 1). This is necessary so that the CVs can be compared across sampling schemes for a given expected number of sets. It is not appropriate to compare them for a given proportion p as the associated expected number of sets sampled is different across the three sampling schemes. This occurs because of rounding and requiring a minimum sample size of 1.

For example, if p = 0.1 and sets are sampled, then every trip that has 1-14 sets will have exactly 1 set sampled (because of rounding and requiring a minimum of 1 set being sampled). The proportion of sets sampled over all trips will be different from 10% as trips with only 1 set will have 100% of their sets sampled, trips with 2 sets will have 50% of their sets sampled, etc. In the case of sampling trips, if p = 0.1, then the total proportion of sets sampled will be close to 10% as there will be less rounding with the number of trips because there will be plenty of them. In the case of sampling trips within vessel there can be a large proportion of sets sampled because, for example, a vessel that has only 1 trip will have that trip sampled and hence all of the sets for that trip will be sampled. The rounding effect is especially important for the observer data as approximately 50% of the vessels have only 1 trip (see Figure 1).

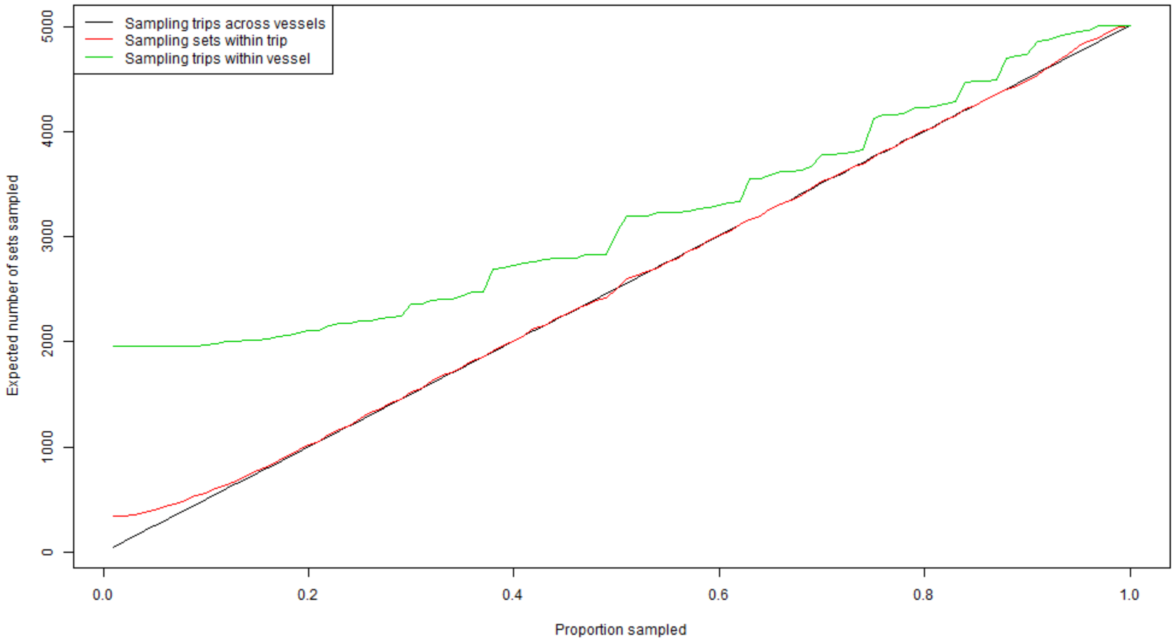

Figure 1: The expected number of sets sampled for each sampling proportion and sampling scheme for observer longline data in 2016 between latitudes 30° S to 10° S.

Development of operating models

The longline logbook data contains information which can be used as a proxy for the full longline fleet when it is eventually largely covered by EM. In each year, for a given area, we have the number of vessels which fished, the trips they did, and the number of sets performed on each trip. There is information on the position of each set, the date and time of each set, the number of hooks set, the number of hooks between floats, and the catch. For species where the catch is not reliably reported, we need to use the limited observer data to estimate the average catch-per-set for each trip and the variance of catch-per-set for each trip. We also need to "tune" the operating model to the variance of catch-per-trip in the observer data (without changing the mean catch-per-trip in the operating model).

The generation of each operating model used three other models: a binomial model to explain the proportion of positive trip catches; a log-linear model to explain mean catch-per-set for the trips that had a positive catch; and a log-linear model to explain the variance of catch-per-set on each trip that had a positive catch.

The models were implemented in R. When there were at least 50 trips, with a positive catch for the given species in the given area, the models were:

```
modbin <- glm(pos ~ days+year+month+lat+long+hkbf+hooks, data=df, family=binomial)

modmean <- lm(log(mean) ~ days+year+month+lat+long+hkbf+hooks, data=df)

modvar <- lm(log(var) ~ log(mean)+days+year+month+lat+long+hkbf+hooks,data=df)
```

Where "pos" is an indicator variable for a positive trip catch, "mean" is the mean catch-per-set for positive trip catches, "var" the corresponding variance, and "days", "lat", "long", "hkbf", and "hooks" are all factors (generally with 8 levels each with a similar number of elements in each level). "year" was a factor with levels for each year. "month" had a factor for each month when there were enough positive catch trips. There was a requirement for at least 30 elements in each level for each model.

"df" is the dataframe that has the observer data for each trip for the given species and area. It contains the various variables before they are turned into factors. "days" is the trip duration. "year" and "month" were the modes for the dates of the sets during each trip; "lat", "long", "hkbf", and "hooks" were the median values for the sets during each trip being respectively latitude, longitude, hooks between floats, and hooks set. Note that the mean is used to explain the variance and a power relationship is assumed (this works well as the variance is largely determined by the magnitude of the mean).

When there were less than 50 trips with positive catches then simpler models were used for the mean and variance:

```
modmean <- lm(log(mean) ~ year, data=df)

modvar <- lm(log(var) ~ log(mean), data=df).
```

The fitted models were used to fill in the mean and variance for each logbook trip (using the specific factors for each trip). In summary, the following steps were taken. Using the binomial model, the probability of obtaining a positive trip catch was calculated for each trip and, using that probability, a value of 0 or 1 was randomly generated. For the trips with positive catches, the mean model was used to calculate a predicted mean catch-per-set for each trip. The mean catches were then tuned so that the trip catches had a variance that matched the variance of the trip catches in the observer data (see Appendix 2). The tuning was done by adding a $N(0, \sigma^2)$ random variable to each model prediction (which is in log space) and adjusting σ^2 until the operating model trip variance was within 10% of the observer trip variance (see Appendix 2). Finally, the variance of the set catches for each trip were predicted using the variance model and adding a random residual (from the model fit).

The operating model approach was applied to ALB, BET, BUM, SKJ, and YFT in each of the two areas 30°-10° S and 10° S to 20° N. It was also applied to SKJ in some sub-areas as there was an issue with the results for 10° S to 20° N. Finally, it was applied to LKV in both areas but these results were not checked against the logbook data as this species is almost never reported in the logbooks.

It should be noted that the operating models are stochastic and that two different operating models for the same species and area are different in the "detail". For target species, there is some variability in predicted total catch, for example, but little variability in the predicted precision of the various EM sampling schemes. For species that are rarely caught there could be considerable variation in the predicted total catch and hence variability in the predicted precision of different EM sampling schemes. In the final analyses, with the chosen species and areas, 200 operating models were generated and the median estimates of predicted precision were used (see below).

For each area and each species (except LKV), a plot was produced comparing the estimated precision under the three different sampling schemes for the operating model and the actual logbook data. For LKV, the plots of estimated CV for the three sampling methods ("sets","trips", and "trips within vessel") were produced.

Analysis of alternative sampling schemes and stratification

An analysis of different sampling schemes was undertaken with a limited set of species to see if any methods looked like they could perform as well as the simple approach of sampling a proportion of sets from all trips.

One general approach is to randomly select: vessels, trips, and sets. In this approach the estimator of total catch (or hooks or any other field for which a total is needed) is built on the average catch-per-set for the selected vessels and trips. For each selected trip, the average catch-per-set, for the selected sets, is scaled by the number of sets to give an estimator of the trip catch. For each selected vessel, the average trip catches are scaled by the number of trips to give an estimator of vessel catch. Finally, the average vessel catch is scaled by the number of vessels to give the estimator of total catch (see Appendix 3).

A second general approach is to ignore vessels and make a random selection of trips and sets. The estimator of total catch is again built on the average catch-per-set, of the selected sets, for each selected trip. The average catch per set is scaled by the number of sets in each trip to give an estimator of trip catch. The average trip catches are then scaled by the number of trips to give an estimator of total catch (see Appendix 3).

The sampling methods are applied to a given species, a given area, over a given timeframe. This captures a particular group of vessels, trips, and sets. These can be partitioned in any number of ways into strata with the hope of minimising the overall variance of the estimator of total catch. For example, a large proportion of the between-trip variation in catch will be due to the number of sets within each trip (e.g, a trip with 100 sets will be expected to catch more than a trip with 10 sets).

The effect of different types of stratification were investigated to see if there were options that would perform better than simply sampling a given proportion of sets for all vessels and trips. These included: stratification of vessels by flag or by the total number of hooks set (during the time period); stratification of trips using the number of sets per trip; and stratification of vessels and trips. Alternative allocations were also explored with different sampling efforts being used within different strata.

Determination of appropriate sampling proportion

It became clear from the preliminary analyses that sampling a given proportion of sets for all vessels and trips would almost always outperform any other sampling scheme. Therefore, the final analyses were restricted to this single approach.

Two main areas were used in the analyses: 30°-10° south; and 10° south to 20° north. The main timeframe considered was a calendar year. Data from sets within the years 2016-2019 inclusive were used. The focus was on total catch with secondary consideration of CPUE (total catch divided by total hooks sets; or total catch divided by total sets)

The species to consider were chosen separately for longline and purse seine without considering the area. The observer data over the whole timeframe (sets within the calendar years 2016-2019 inclusive) were used to calculate the CV of catch per set for each species (approximately 78,000 longline and 119,000 purse seine sets). For the longline data, species with fewer than 100 occurrences in the catch were not considered (this left 110 out of 357 species). For the purse seine data, species with fewer than 30 non-zero catches were not considered (this left 76 out of 220 species). The species with the highest and lowest CV of catch per set in the categories: billfish, birds, marine mammals, sharks and rays, marine turtles, and tuna; were automatically chosen. Additional tuna species were included as needed so that ALB, BET, SKJ, and YFT were all included. Additional species were included to fill in the range of the CV of catch per set to bring the total number of species to 15 for each method (see Table 1).

For a given sampling proportion, the precision of a field total (i.e., total catch or total effort) will depend on the number of sets involved in the total. The larger the number of sets the more precise the estimator of the total will be. As we move from a calendar year over one of the main areas to say 6 months over a "regional" area, there will be a reduction in the expected precision of estimators. The "shape" of this curve was investigated by looking at totals for each of the flag states and also by looking at systematic reductions in the size of an area and the timeframe.

The field, hooks between floats (hkbf) was considered for the longline data. The current policy for observers and EM analysts is to count the hkbf for the first 3 baskets, the "middle" 3 baskets and the last 3 baskets and average them to give a single whole number for each observed set. The usefulness of this approach was considered for an EM analyst.

Table 1: The species to be used in the final study for each fishing method together with the CV of catch per set in the observer data (being the highest and lowest for each species category).

	Longline			Purse seine		
	Code	**Name**	**CV**	**Code**	**Name**	**CV**
Billfish	BUM	Blue marlin	2.5	BUM	Blue marlin	6.7
	SPF	Long-billed spearfish	20.2	SSP	Short-billed spearfish	75.6
Birds	DKN	Black footed albatross	12.9	-		
	TWD	White capped albatross	50.0	-		
Mammals	SEA	NZ fur seal	46.8	FAW	False killer whale	51.2
	-			SHW	Short-finn pilot whale	110.0
Sharks & rays	PLS	Pelagic stingray	2.2	FAL	Silky shark	6.2
	LMD	Salmon shark	91.1	ALS	Silvertip shark	124.1
Turtles	LKV	Olive ridley turtle	13.0	LKV	Olive ridely turtle	40.5
	DKK	Leatherback turtle	25.7	TTH	Hawksbill turtle	190.5
Tuna	BET	Bigeye	1.4	SKJ	Skipjack	1.3
	SLT	Slender	48.2	ALB	Albacore	75.2
Target tuna	ALB	Albacore	1.8	BET	Bigeye	4.0
	YFT	Yellowfin	1.9	YFT	Yellowfin	2.0
	SKJ	Skipjack	3.3			
Additional species	PBF	Pacific bluefin tuna	33.0	RHN	Whale shark	21.4
				BLR	Blacktip reef shark	95.6
				CXS	Bigeye trevally	133.0

Regional considerations

Consideration was also given to how a manager in a regional centre might decide what sampling proportion was appropriate for their region given the size of their fleet and their species of concern. An approximate formula for the CV of a catch total was developed for species which are caught infrequently (i.e., 5-20% of the sets) and the distribution of the total catch estimator was approximated with a lognormal distribution. Also, consideration was given to the required sampling proportion to check whether a particular species had been caught or not (i.e., of relevance to species of special interest).

An approximate formula for the CV of a catch total when sampling a proportion (*p*) of the sets for each trip comes from noting that this is a stratified version of sampling all sets (across all trips) at random. For infrequently caught species there will be only a small benefit in stratifying by trip and the CVs will be similar.

Let the catches on each of the *N* sets be $\{ c_1, \ldots, c_N \}$ and assume random sampling without replacement with sample size $n = pN$. Denote the *i*th sample as X_i and define the total catch estimator as

$$T = N\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)$$

It follows that *T* is an unbiased estimator of the total catch and

$$CV(T) \approx \sqrt{\frac{1-p}{p}}\frac{1}{\sqrt{N}}\frac{\sigma}{\mu}$$

where μ and σ are respectively the population mean and standard deviation of the catches. The approximation only assumes that $1/N$ is small (e.g., $N > 100$).

When catches are infrequent, the formula can be further simplified by assuming that the positive catches are all equal. This is a good approximation because the bulk of the population variance will come from the deviations from the mean at the zero catches.

The formula for the CV then becomes

$$CV(T) \approx \sqrt{\frac{1-p}{p}}\frac{1}{\sqrt{N}}\sqrt{\frac{1-r}{r}}$$

where *r* is the proportion of positive catches. Note, that the CV does not depend on the value of the positive catches. This formula was used with the assumption of a lognormal distribution to approximate the distribution of the total catch estimator at different values of *p*, *r*, and *N* (for infrequent catches). The lognormal assumption is not valid when $Nr < 5$ as in those cases there are not enough supports in the discrete domain to allow a good approximation by a continuous distribution (e.g., $N = 100$ and $r = 0.01$ has only two possible outcomes for the total catch estimator for each given sampling proportion: 0 or whatever the scaled single positive catch is).

For some species of special interest, the main question may be whether any of that species has been caught or not. In this case, it is important to sample enough sets so that there is high confidence that if an animal was caught then a positive total catch estimate will be observed. Under the assumptions above, the number of positive catches observed follows a hypergeometric distribution (which is the "binomial distribution" for when sampling without replacement). For different values of *r* and *N*, the sampling proportion (*p*) required to reduce the probability of falsely observing a zero catch to a low level was calculated and plotted.

Results

Preliminary analyses

One of the important points when considering results is that the CV of an estimator for a total of *n* elements will generally decrease as *n* increases. A special case of this is the sum of *n* independent random variables with the same variance and mean. In this case the CV is exactly proportional to 1/sqrt(*n*). In the more general case, provided that the random variables are independent and have similar mean and variance, the CV of the total will be approximately proportional to 1/sqrt(*n*).

Consequently, the precision of an estimator of a total will depend on how many vessels, trips, and sets are involved. Generally, the more there are of each, the smaller the CV of the estimator will be. So, for example, the CVs for each value of *p* when sampling sets for total ALB catch, in a pooled group of years, will be lower than the CVs for the individual years (Figure 2).

This issue is also relevant when considering results from the observer data compared to results from the logbook data. The logbook data consists of about 7 times as many trips for any given year and it is to be expected that the CVs of estimators of totals will be much lower for the logbook data than for the observer data (e.g., Figure 3). Therefore, it is crucial to consider results generated for the "whole fleet" and not to draw conclusions, in any absolute sense, from results obtained from just the observer data.

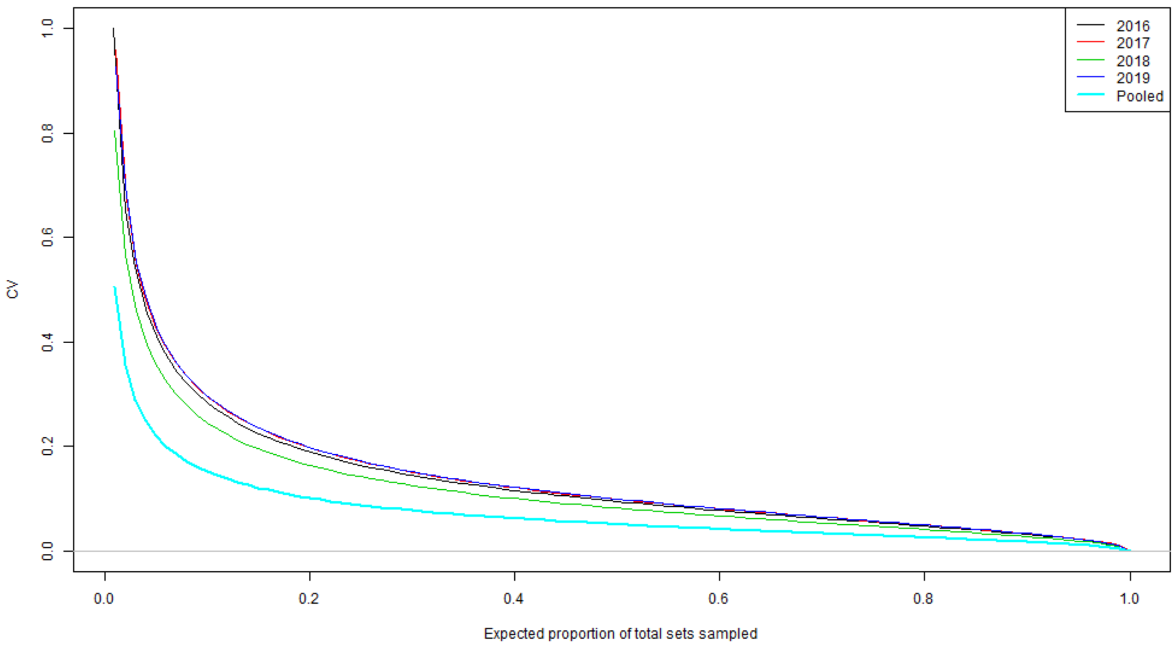

Figure 2: The CV of the "sampling sets" estimator of total ALB catch from 30° S to 10° S in the years 2016-2019 individually and when pooled for the observer data. The CVs were calculated for each proportion from 0.01 to 1 in steps of 0.01 and are plotted against the associated expected proportion of total sets sampled.

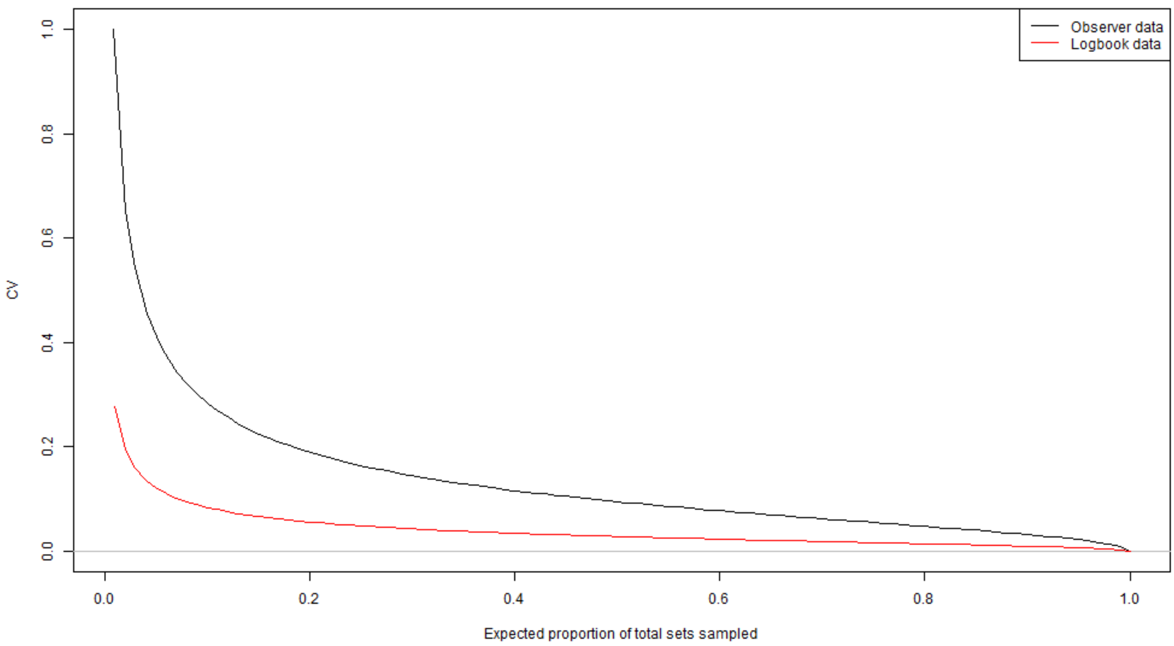

Figure 3: The CV of the "sampling sets" estimator of total ALB catch from 30° S to 10° S in 2016 for the observer data and the logbook data. The CVs were calculated for each proportion from 0.01 to 1 in steps of 0.01 and are plotted against the associated expected proportion of total sets sampled.

The CVs of the estimators of total catch will also depend on the average catch rates of the species and the variance of catches between sets, between trips, and between vessels. The logbook data is unreliable for catches except for the main target species. For bycatch species, the observer data must be used but it must be "scaled up" to the full fleet otherwise the results will not be useful in an absolute sense (i.e., in terms of the absolute values of the CVs).

Some of the preliminary results are presented below, for total catch and total hooks, where the three different sampling schemes are compared for observer data and logbook data (main target species). The results are very similar to those obtained by Peatman & Nicol (2020) for nominal CPUE with "sampling sets" being the best method as it avoids any between-trip or between-vessel variance.

When sampling sets, the variance of hooks set is tiny compared to the total hooks set in any given year for the areas considered in these examples (30° S to 10° S and 10° S to 20° N) (Figures 4 & 5). For nominal CPUE, the CVs of the estimators will be almost identical to the CVs for total catch (so there is no need to directly estimate the CVs for nominal CPUE). The same result holds for hooks observed where the four years have been pooled to get the total hooks observed on a similar scale to the fleet for a single year (Figure 6).

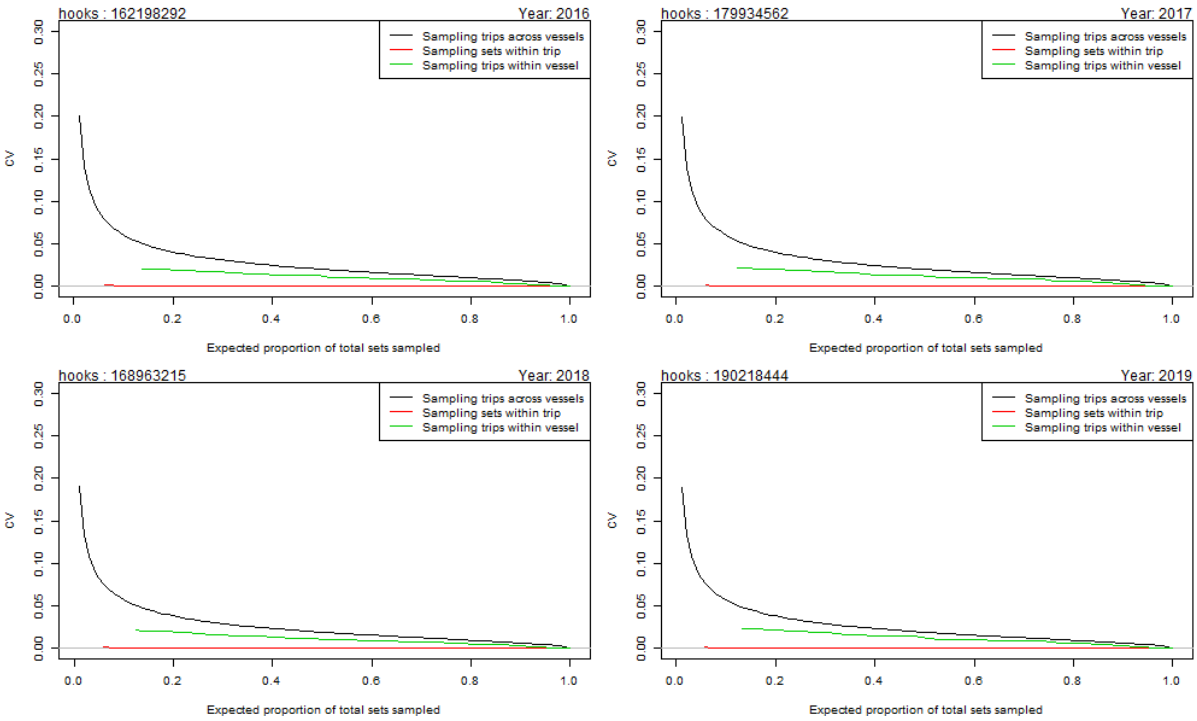

Figure 4: For the three different sampling schemes, the CV of the estimator of total hooks set from 30° S to 10° S in 2016 to 2019 for the logbook data. The CVs were calculated for each proportion from 0.01 to 1 in steps of 0.01 and are plotted against the associated expected proportion of total sets sampled.

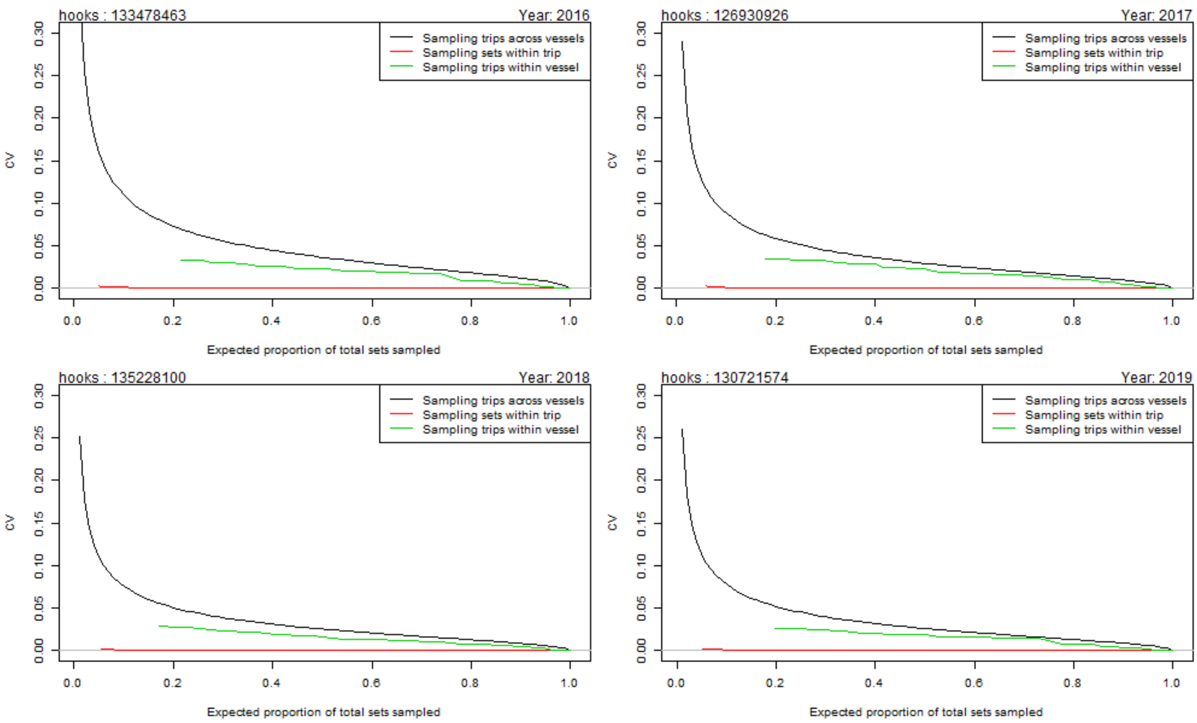

Figure 5: For the three different sampling schemes, the CV of the estimator of total hooks set from 10° S to 20° N in 2016 to 2019 for the logbook data. The CVs were calculated for each proportion from 0.01 to 1 in steps of 0.01 and are plotted against the associated expected proportion of total sets sampled.

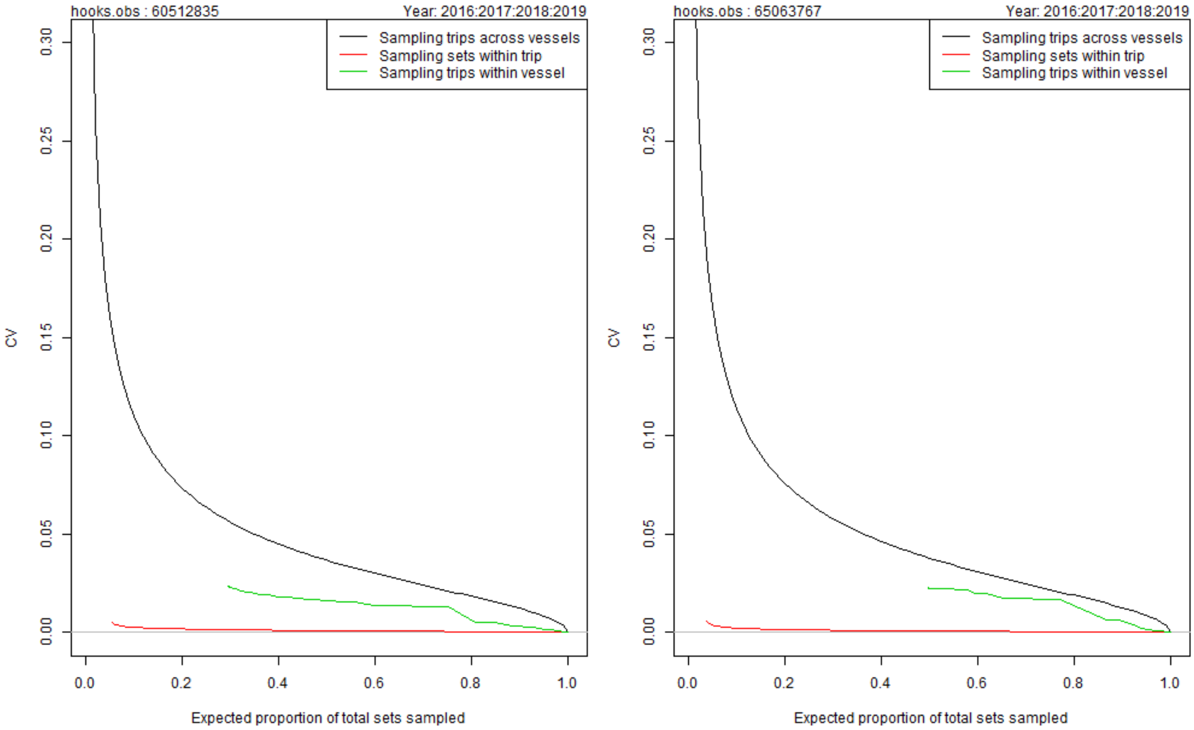

Figure 6: For the three different sampling schemes, the CV of the estimator of total hooks observed from 30° S to 10° S (left) and from 10° S to 20° N (right) in 2016 to 2019 (pooled) for the observer data. The CVs were calculated for each proportion from 0.01 to 1 in steps of 0.01 and are plotted against the associated expected proportion of total sets sampled.

The result that sampling for sets provides the most precise estimators of the three sampling schemes holds for species catch number totals as well as total hooks. For ALB, YFN, and BET, sampling for sets provides very precise estimators at all values of *p* (even 0.01, although this corresponds to about 5-6% of sets being sampled) (Figures 7-9).

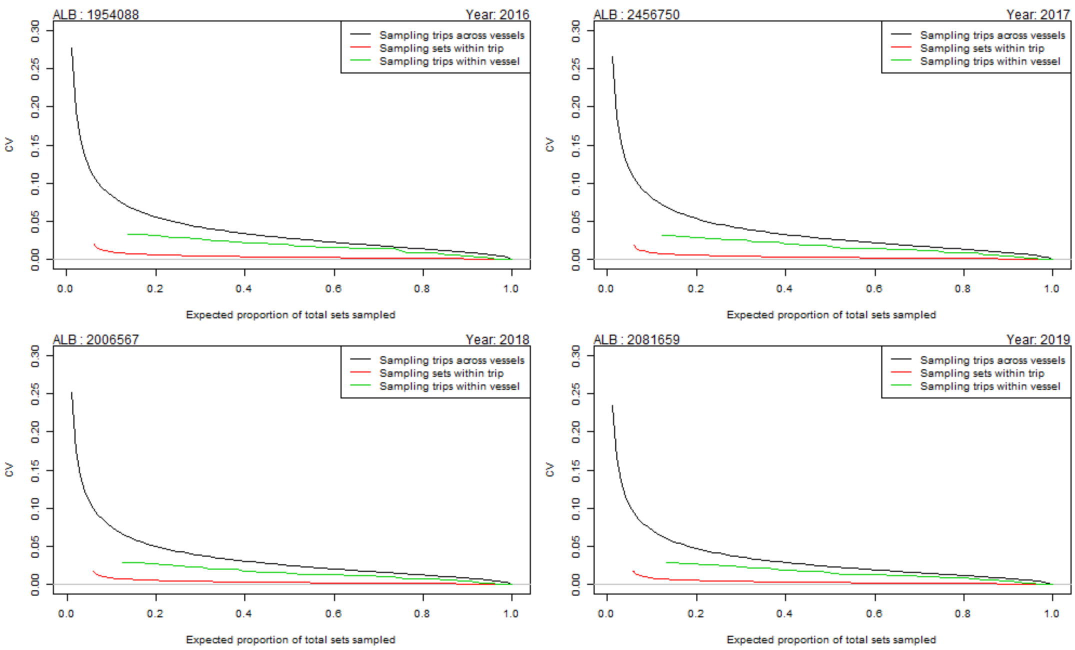

Figure 7: For the three different sampling schemes, the CV of the estimator of total ALB catch from 30° S to 10° S in 2016 to 2019 for the logbook data. The CVs were calculated for each proportion from 0.01 to 1 in steps of 0.01 and are plotted against the associated expected proportion of total sets sampled.

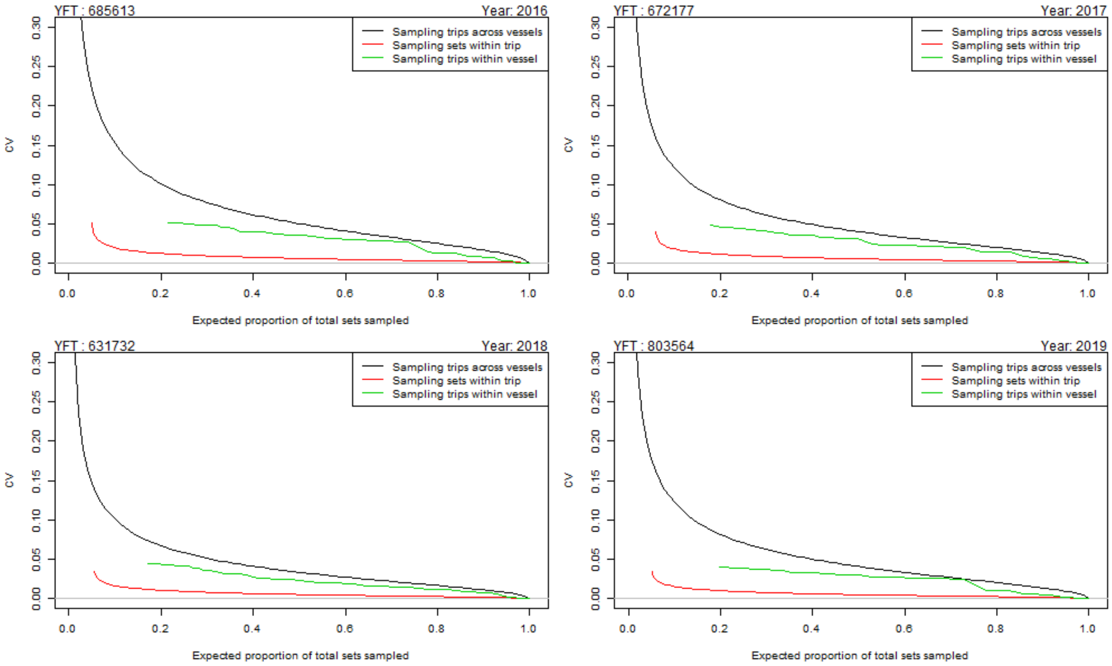

Figure 8: For the three different sampling schemes, the CV of the estimator of total YFT catch from 10° S to 20° N in 2016 to 2019 for the logbook data. The CVs were calculated for each proportion from 0.01 to 1 in steps of 0.01 and are plotted against the associated expected proportion of total sets sampled.

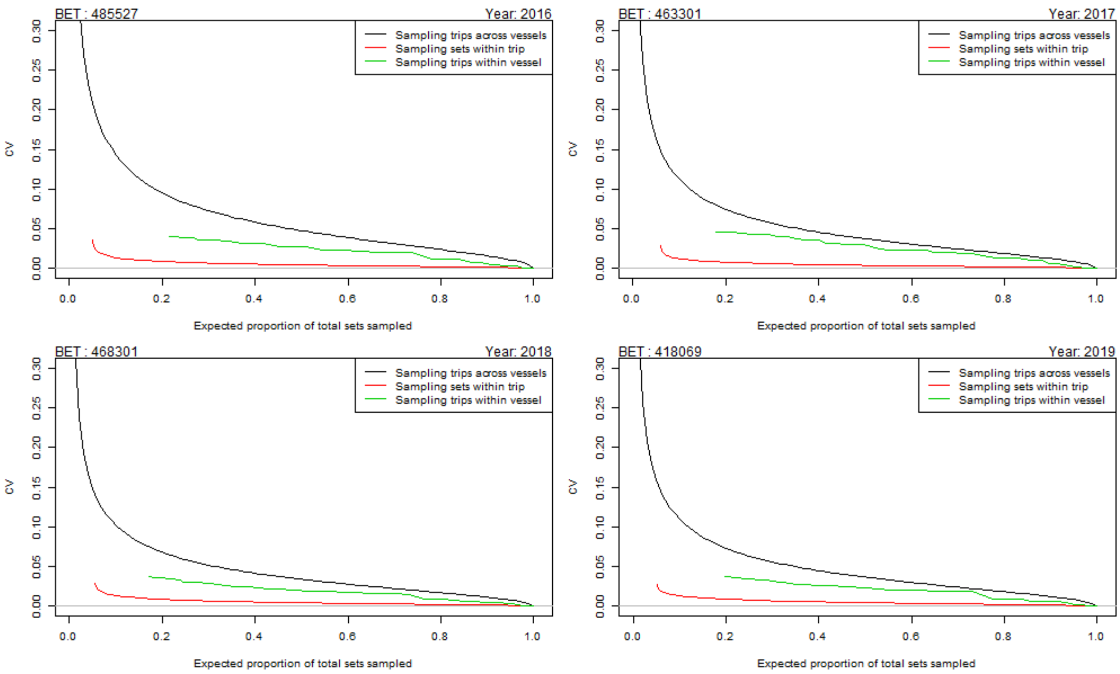

Figure 9: For the three different sampling schemes, the CV of the estimator of total BET catch from 10° S to 20° N in 2016 to 2019 for the logbook data. The CVs were calculated for each proportion from 0.01 to 1 in steps of 0.01 and are plotted against the associated expected proportion of total sets sampled.

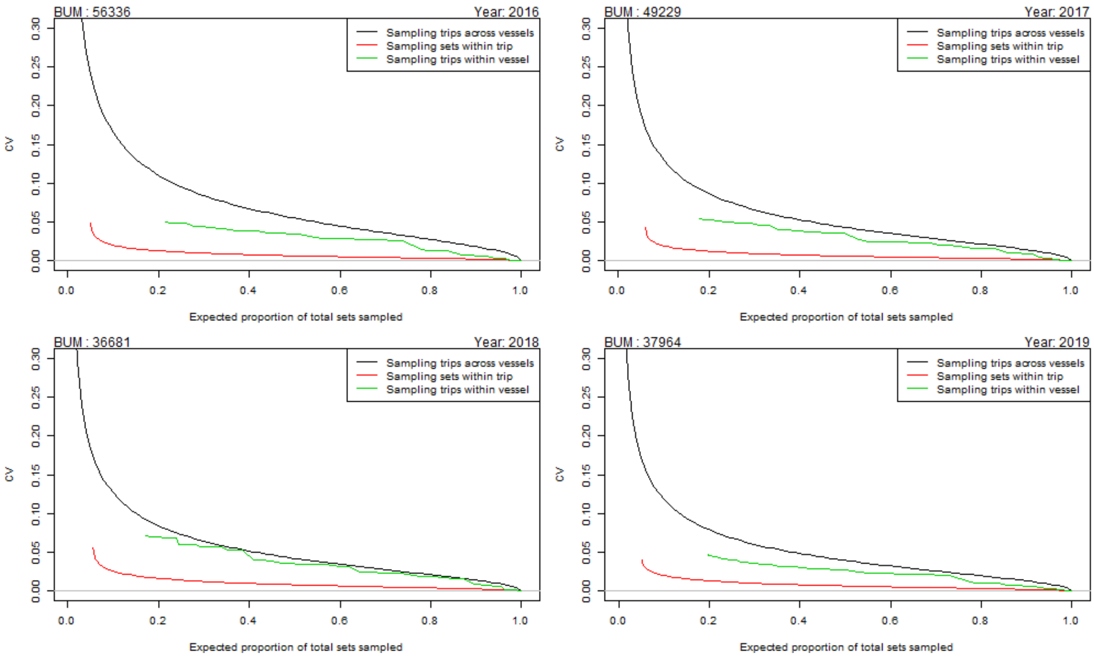

Figure 10: For the three different sampling schemes, the CV of the estimator of total BUM catch from 10° S to 20° N in 2016 to 2019 for the logbook data. The CVs were calculated for each proportion from 0.01 to 1 in steps of 0.01 and are plotted against the associated expected proportion of total sets sampled.

For BUM and SKJ, annual catch numbers are much lower than for ALB, YFT, and BET. This is reflected in the precision of the catch total estimators with higher sampling proportions being needed to achieve CVs less than 5% when sampling sets (Figures 10 & 11). For LKV, which is rarely caught, there is still a difference in the CVs achieved by the three sampling methods but they are all above 5% until about 60% of the sets are sampled (Figure 12).

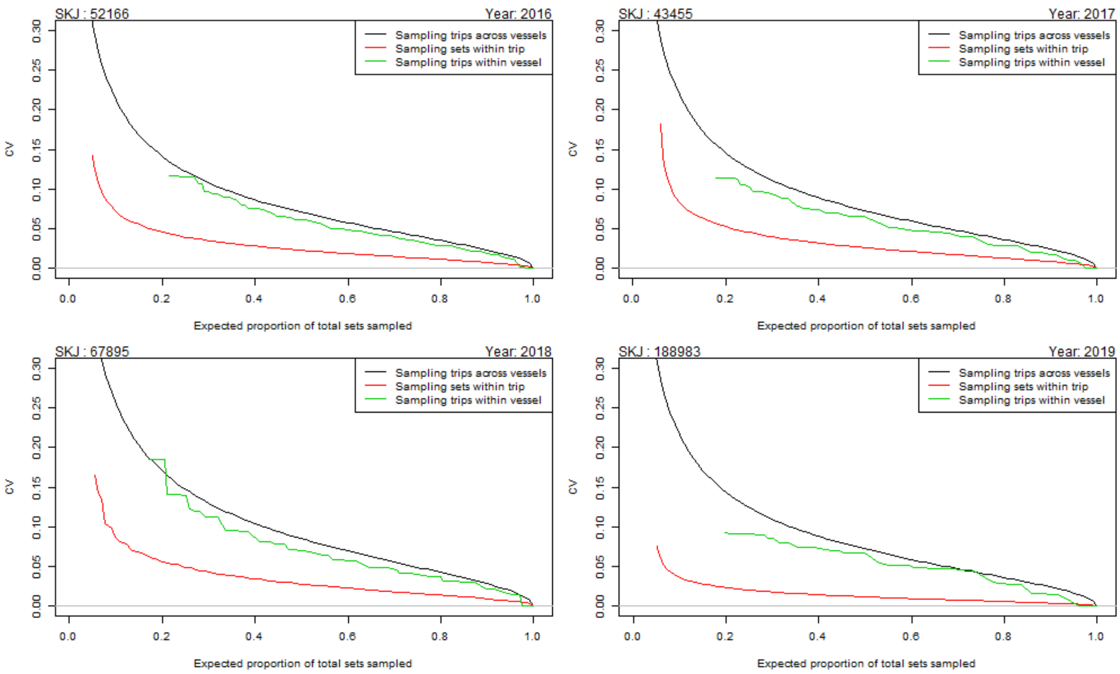

Figure 11: For the three different sampling schemes, the CV of the estimator of total SKJ catch from 10° S to 20° N in 2016 to 2019 for the logbook data. The CVs were calculated for each proportion from 0.01 to 1 in steps of 0.01 and are plotted against the associated expected proportion of total sets sampled.

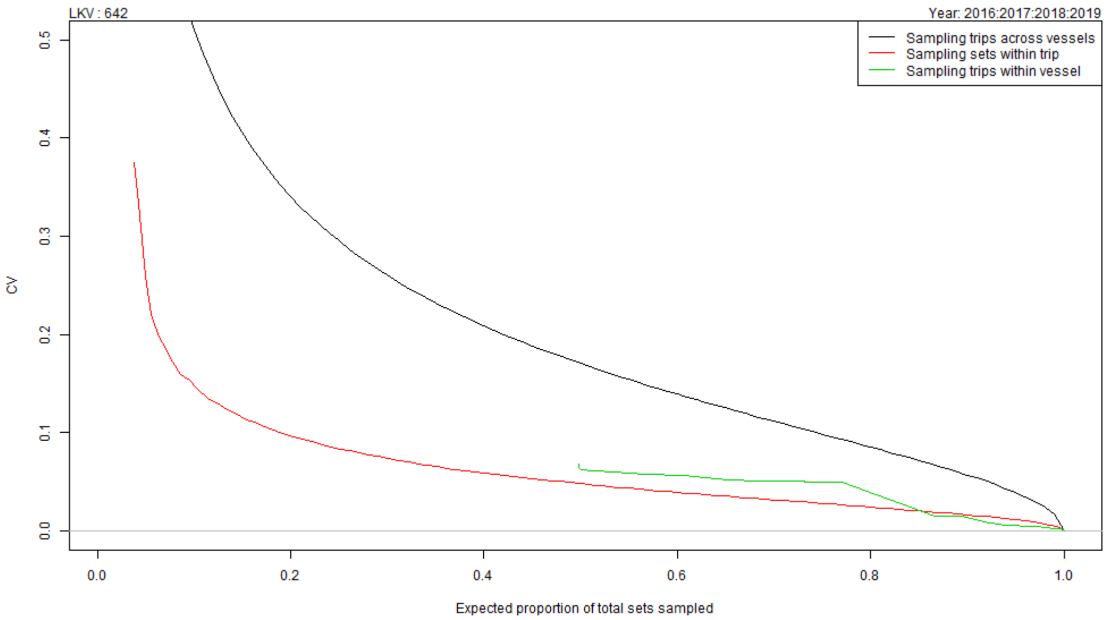

Figure 12: For the three different sampling schemes, the CV of the estimator of total LKV catch from 10° S to 20° N in 2016 to 2019 (pooled) for the observer data. The CVs were calculated for each proportion from 0.01 to 1 in steps of 0.01 and are plotted against the associated expected proportion of total sets sampled.

Development of operating models

The binomial model fitted the data adequately in that the predicted proportion of positive catches always closely matched the actual proportion of positive catches (Table 2). The trip-duration factor was generally significant as longer trips have less chance of a zero catch. Sometimes other variables were also significant.

For the log-linear models (with the exception of LKV in the southern area) the adjusted R^2 values ranged from 20-59% for the mean models and from 81-96% for the variance models (Table 2). The mean models had a variety of variables showing up as significant depending on the species and area. The variance models always had the mean and duration as highly significant.

Table 2: Measures of the goodness of fit for the binomial and log-linear models for each species in the northern and southern areas. "Positive" is the percentage of positive trip catches in the observer data; "Ratio" is the expected number of positive trip catches from the binomial model divided by the actual number of positive trip catches; the R^2 values are adjusted R^2 for the mean and variance models. Note, for LKV in the southern area, where there were less than 50 trips with a positive catch, only year was used as an explanatory variable in the mean model and only the mean was used as an explanatory variable in the variance model.

	ALB	BET	BUM	SKJ	YFT	LKV
Northern						
Positive (%)	63	98	81	87	98	16
Ratio (%)	99	98	99	99	99	99
R^2 (mean)	42	40	20	40	39	34
R^2 (var.)	96	81	85	89	89	96
Southern						
Positive (%)	97	93	57	86	99	3
Ratio (%)	100	100	100	99	100	100
R^2 (mean)	59	35	25	22	31	3
R^2 (var.)	88	87	88	91	83	100

For LKV, there were only 37 positive observer trip catches in the southern area (with 188 positive ones in the northern area). Therefore, the simpler log-linear models were fitted for the southern area and the R^2 value was only 3% for the mean model (i.e., year did not explain much of the variation in mean catch-per-set within trip). The variance model had a very high R^2 showing that the variance was almost fully explained by the mean (Table 2). The prediction of CV when sampling using "sets" will be accurate for this reason (i.e., because the estimates of CV depend on the relationship between the mean and variance rather than the absolute value of the mean catch-per-set which is poorly estimated). The trip-based estimates of CV could be very poor.

The main result presented here is the plot of the predicted precision from a realisation of the operating model compared with the precision for the genuine logbook data.

With one exception, the results were good for the 10 combinations of species and area (Figures 13-22). The exception was SKJ in the northern area (Figure 19). In this case the operating model was poor at predicting the precision of the sampling schemes except in 2019. An examination of the positions of the observed sets and the logbook sets showed that from 13°-20° N there is little overlap between the observer and logbook positions. Also, the proportion of positive trip catches (of SKJ) in the observer data was 93% for this area but in the logbook data it was just 32%. It appears that there was substantial observer effort in a fleet that typically catches SKJ but does not fill in logbooks (in an area south of Hawaii).

For rarely caught species, it will be difficult to get a good idea of mean catch-per-set across the logbook fleet. However, the strong relationship between mean and variance of catch-per-set suggests that the operating model should provide a good proxy for the CV that will be achieved by sampling sets. For the single LKV operating model, the estimated precision suggests that sampling 20% of the sets for each trip will give a CV of about 10% on total catch each year in the southern area (Figure 23) and that sampling only 10% of sets would be needed for a CV of about 10% in the northern area (Figure 24).

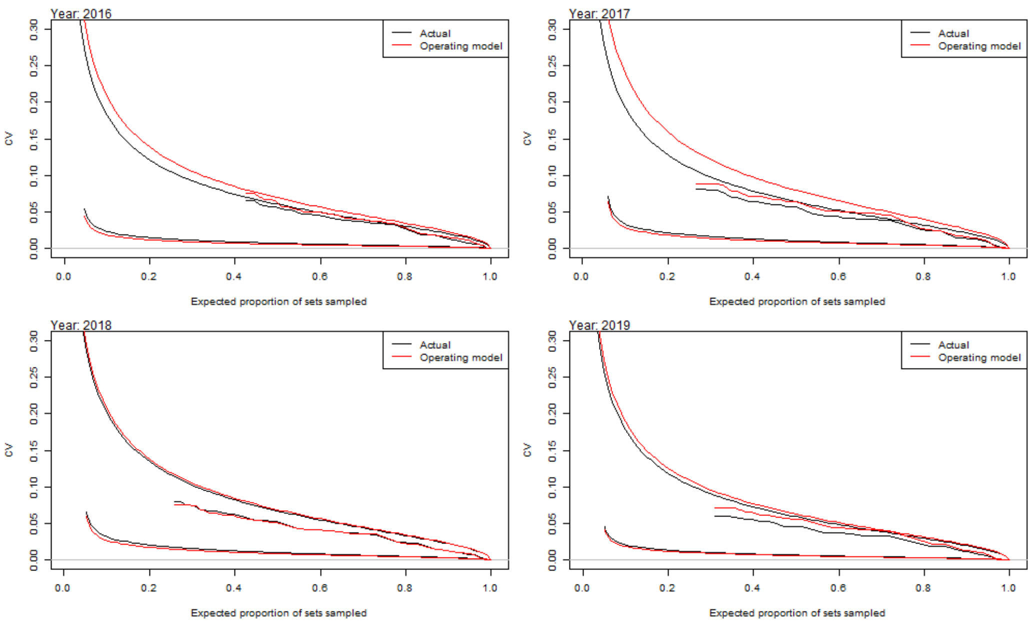

Figure 13: ALB 10° S to 20° N. The estimated precision of the three different sampling schemes for different levels of expected proportion of sets sampled for 2016-2019. Black lines show the results for the logbook data and red lines for an operating model generated from the observer data.

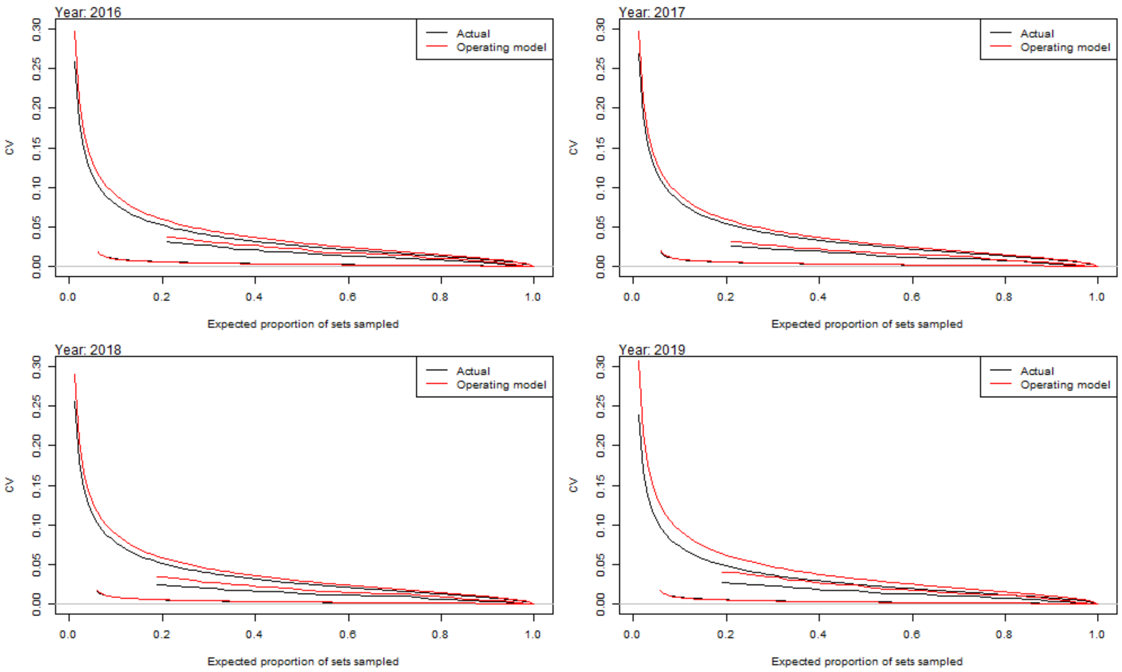

Figure 14: ALB 30° S to 10° S. The estimated precision of the three different sampling schemes for different levels of expected proportion of sets sampled for 2016-2019. Black lines show the results for the logbook data and red lines for an operating model generated from the observer data.

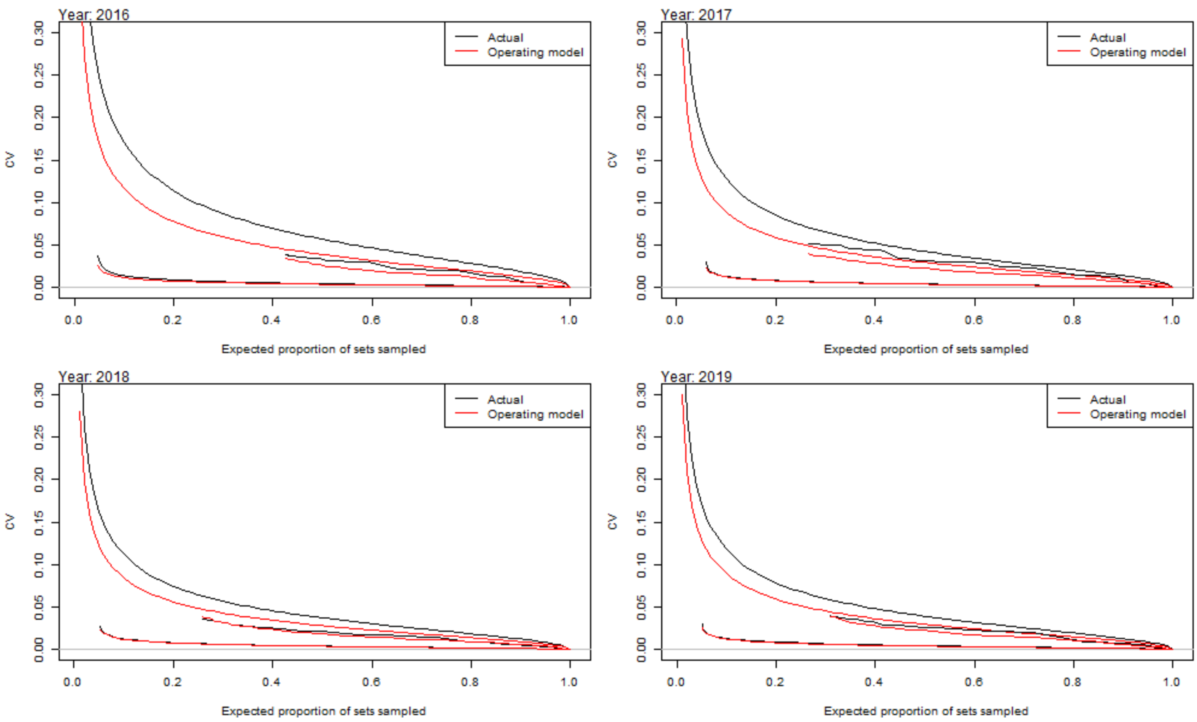

Figure 15: BET 10° S to 20° N. The estimated precision of the three different sampling schemes for different levels of expected proportion of sets sampled for 2016-2019. Black lines show the results for the logbook data and red lines for an operating model generated from the observer data.

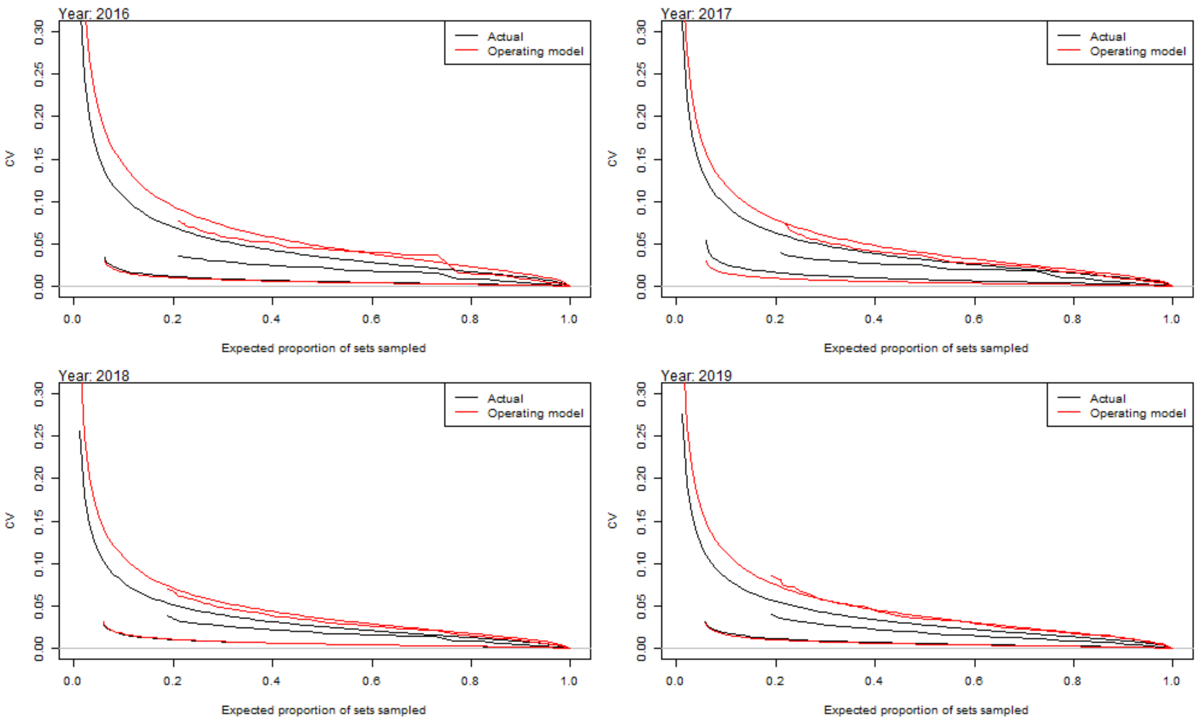

Figure 16: BET 30° S to 10° S. The estimated precision of the three different sampling schemes for different levels of expected proportion of sets sampled for 2016-2019. Black lines show the results for the logbook data and red lines for an operating model generated from the observer data.

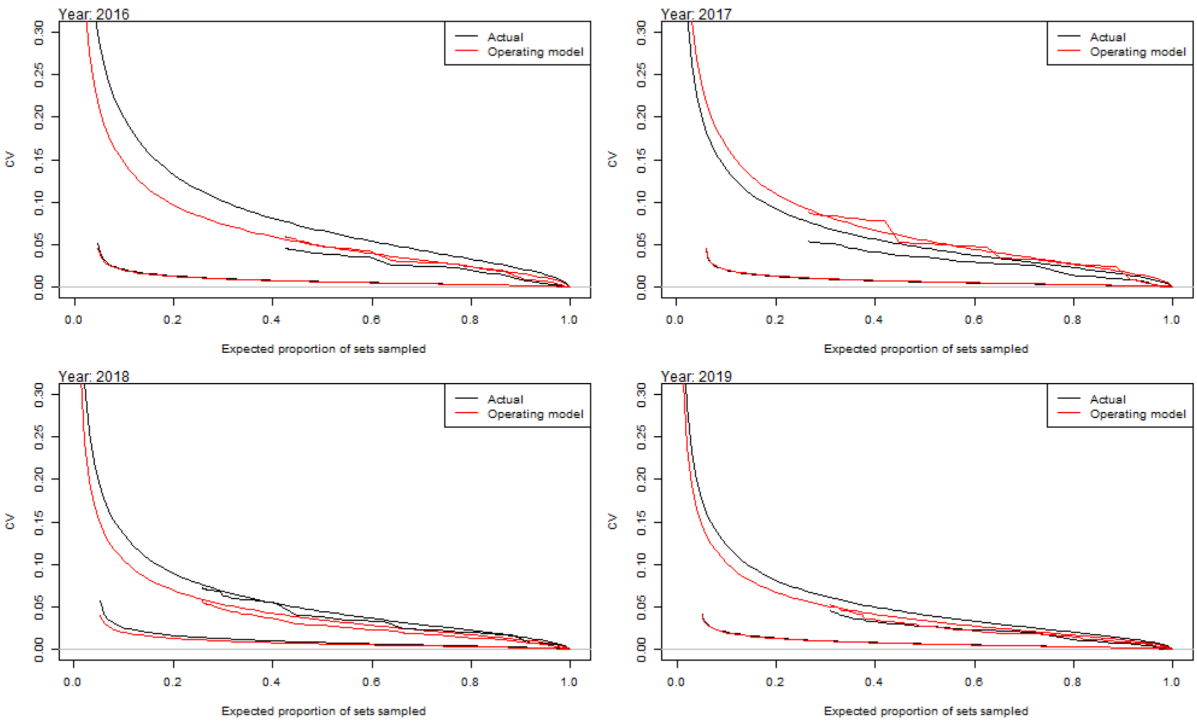

Figure 17: BUM 10° S to 20° N. The estimated precision of the three different sampling schemes for different levels of expected proportion of sets sampled for 2016-2019. Black lines show the results for the logbook data and red lines for an operating model generated from the observer data.

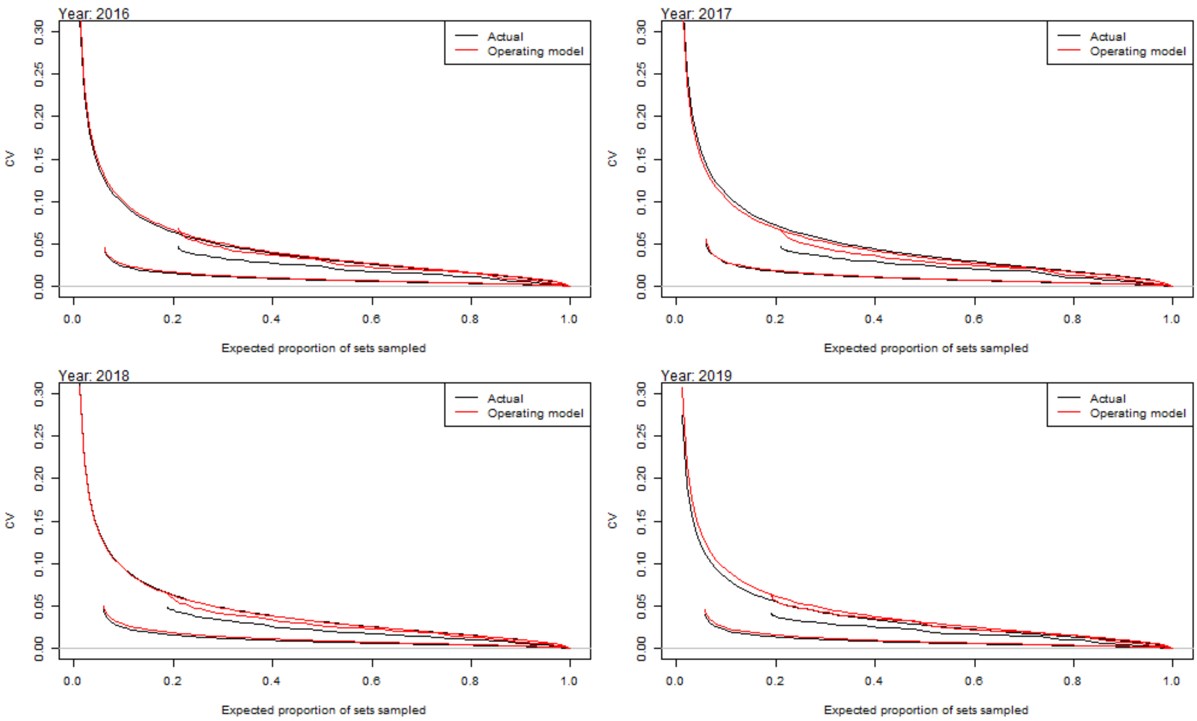

Figure 18: BUM 30° S to 10° S. The estimated precision of the three different sampling schemes for different levels of expected proportion of sets sampled for 2016-2019. Black lines show the results for the logbook data and red lines for an operating model generated from the observer data.

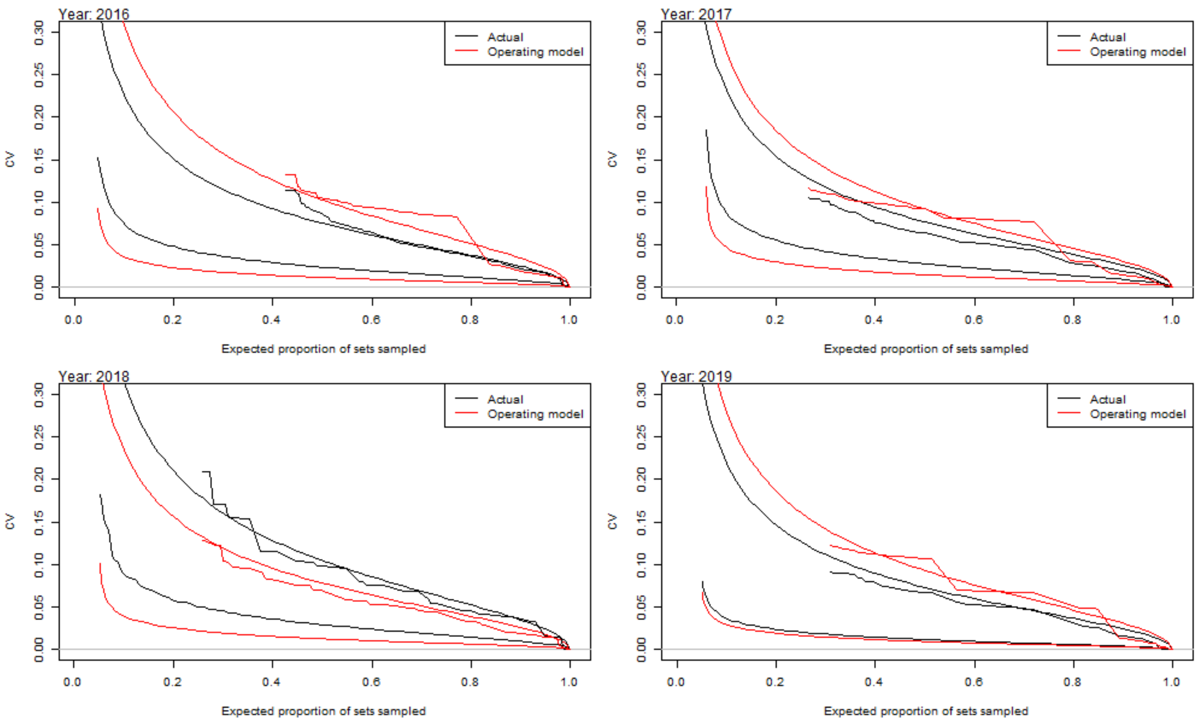

Figure 19: SKJ 10° S to 20° N. The estimated precision of the three different sampling schemes for different levels of expected proportion of sets sampled for 2016-2019. Black lines show the results for the logbook data and red lines for an operating model generated from the observer data.

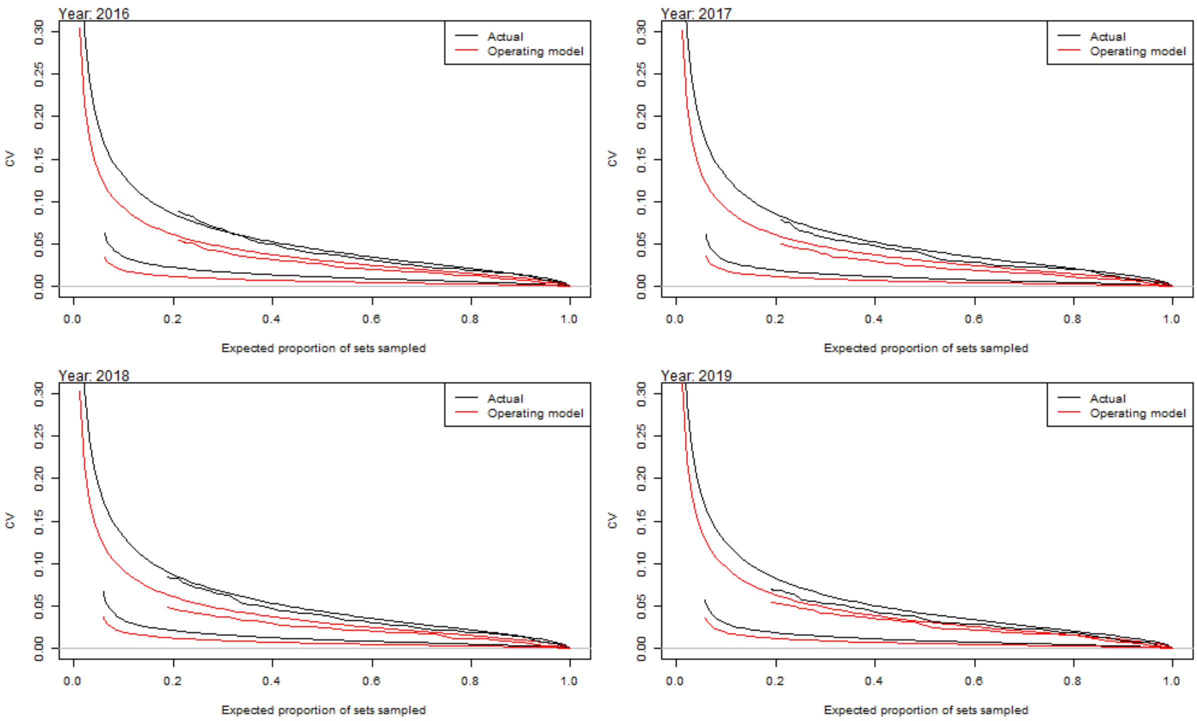

Figure 20: SKJ 30° S to 10° S. The estimated precision of the three different sampling schemes for different levels of expected proportion of sets sampled for 2016-2019. Black lines show the results for the logbook data and red lines for an operating model generated from the observer data.

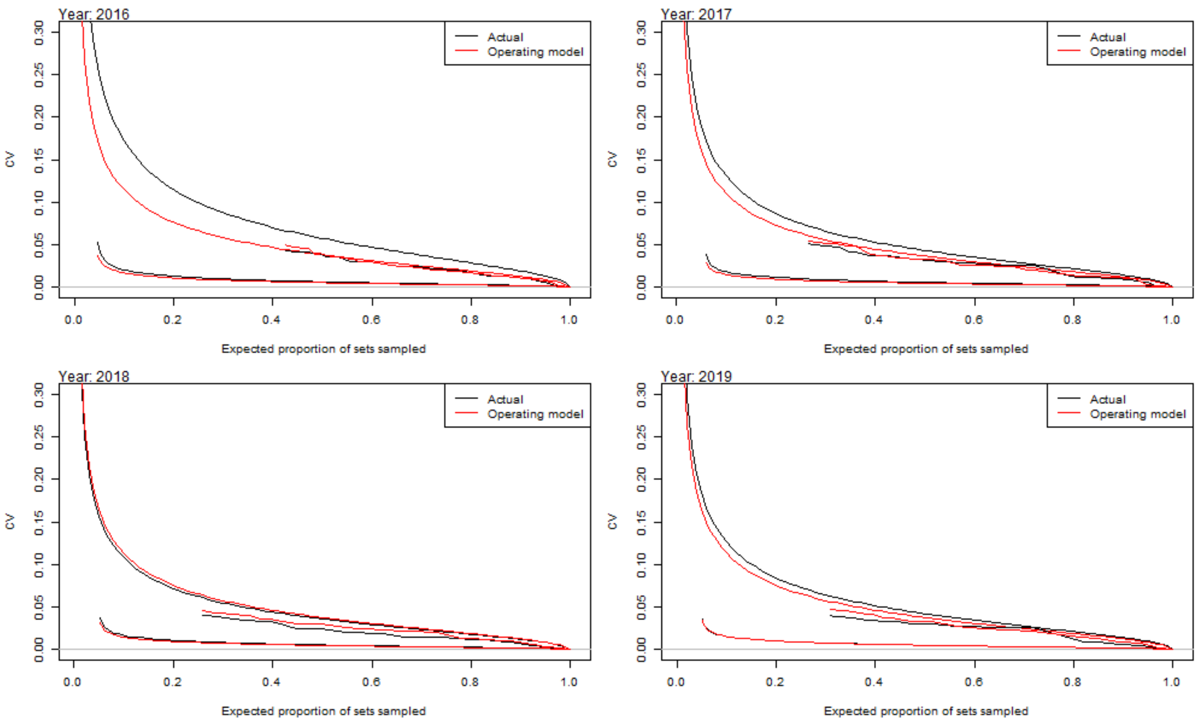

Figure 21: YFT 10° S to 20° N. The estimated precision of the three different sampling schemes for different levels of expected proportion of sets sampled for 2016-2019. Black lines show the results for the logbook data and red lines for an operating model generated from the observer data.

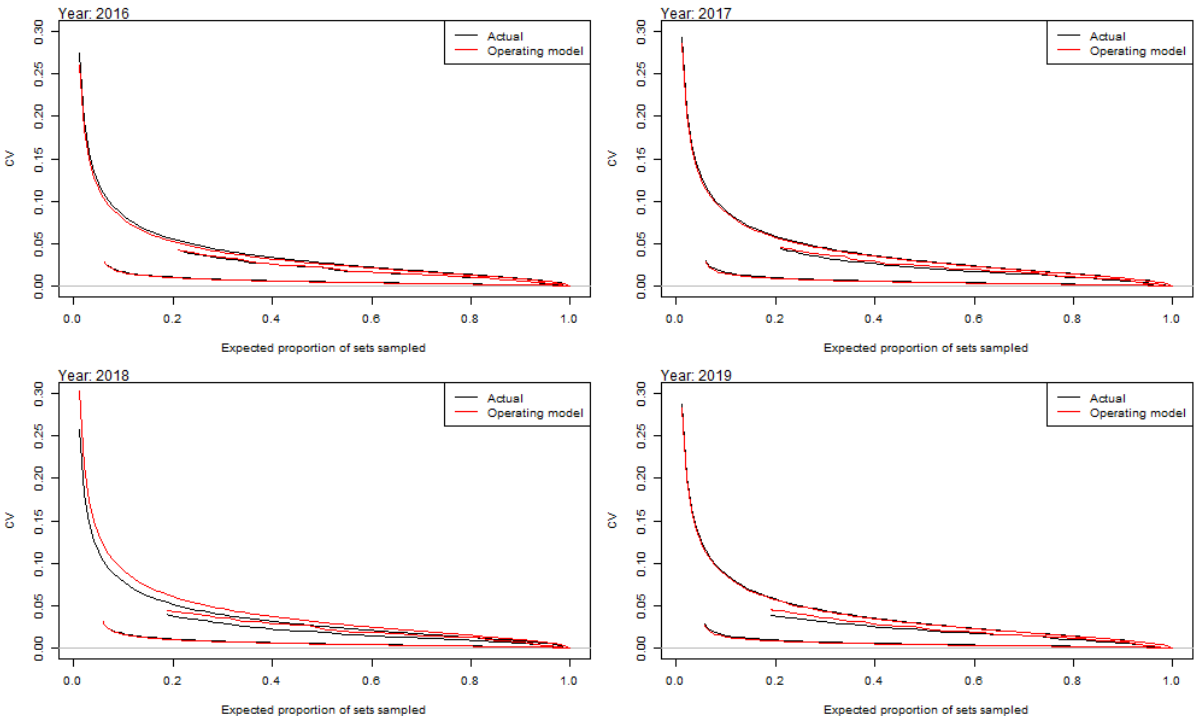

Figure 22: YFT 30° S to 10° S. The estimated precision of the three different sampling schemes for different levels of expected proportion of sets sampled for 2016-2019. Black lines show the results for the logbook data and red lines for an operating model generated from the observer data.

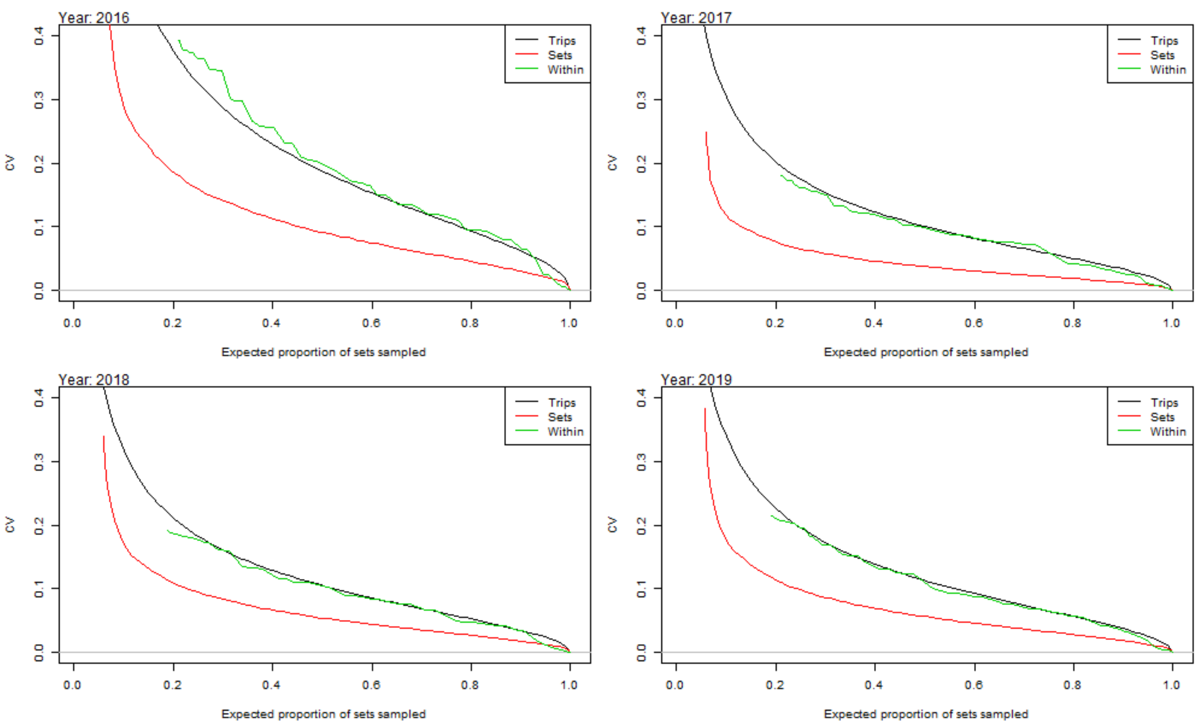

Figure 23: LKV 30° S to 10° S. The estimated precision of the three different sampling schemes for different levels of expected proportion of sets sampled for 2016-2019. The results are from a realisation of the operating model for LKV (a proxy for the full fleet).

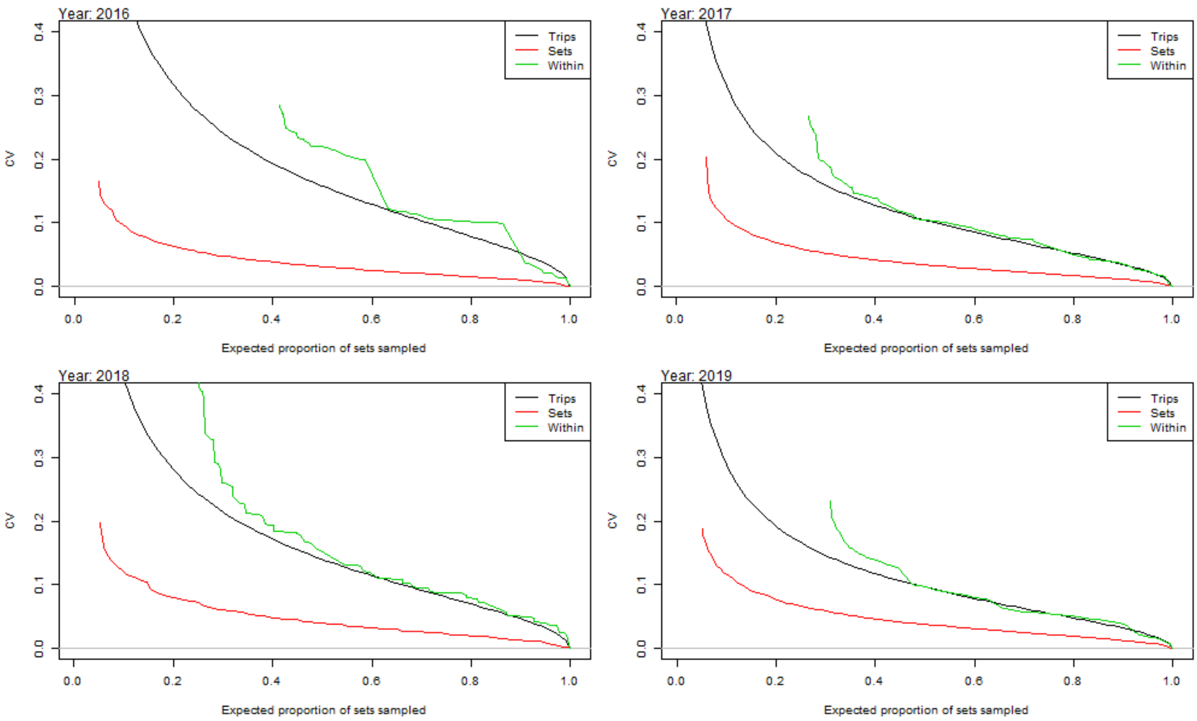

Figure 24: LKV 10° S to 20° N. The estimated precision of the three different sampling schemes for different levels of expected proportion of sets sampled for 2016-2019. The results are from a realisation of the operating model for LKV (a proxy for the full fleet).

Analysis of alternative sampling schemes and stratification

From preliminary analyses it became clear that any sampling scheme that didn't sample all vessels and trips was going to struggle to perform well. This happens because when some trips are not sampled between-trip variation is non-zero and can be very high. Certainly, without stratification, it is obvious that between trip variation can be huge as some trips will have few sets (and relatively little catch) while other trips could have up to about 300 sets. Stratification by trip is therefore essential if there is random sampling of trips in any fashion. Similarly, if not all vessels are sampled then the between-vessel variation can be huge. Therefore, if vessels are sampled at random then there must be stratification by vessels.

However, even with trip and/or vessel stratification, in the examples considered, it was difficult to find better performance than the simple approach of sampling a proportion of sets for all trips. It did happen but only over a very limited range of possible sampling efforts. The examples in this report are all for blue marlin (BUM) but a similar effect was seen for other species.

When sampling for vessel, trip, and sets, with all vessels sampled, a stratification by trip lowered the CVs for four different levels of trip proportion considered (Figure 25). However, the simple approach of sampling a proportion of sets for all trips was still the best (Figure 25). A stratification by trip was also important when considering different levels of vessel selection (Figures 26 & 27).

When vessels were ignored and the sampling scheme was selecting a proportion of sets for a selected proportion of trips, the effect of stratification by trips was still important (Figure 28). Stratification by flag was better than no stratification at all but not nearly as good as stratification by trips (Figure 28).

Stratification by vessel and trip was only a tiny bit better than stratification just by trip when 80% of vessels and 50% of trips were selected (Figure 29). However, when the vessel strata were given different levels of vessel and trip selection (greater sampling effort for vessels that set more hooks in the area during the year) then lower CVs were achieved for the vessel and trip stratification compared to just trip stratification (Figure 30). Indeed, for a tiny part of the curve, the vessel and trip stratification had lower CVs than “sampling sets” (Figure 30).

Of course, given that decisions need to be made during the year when data become available, rather than at the end of year when all trips are completed, stratification by vessel based on the number of hooks set will not be possible. It will be possible to stratify vessels for the coming year on the basis of data collected in previous years or given general knowledge of the size of the vessel and planned activities. The point of the vessel stratification is to show that even with advance knowledge of the number of hooks set the stratification does not lead to better performance than that achieved by the simple approach of sampling all vessels and trips with a constant proportion of sets.

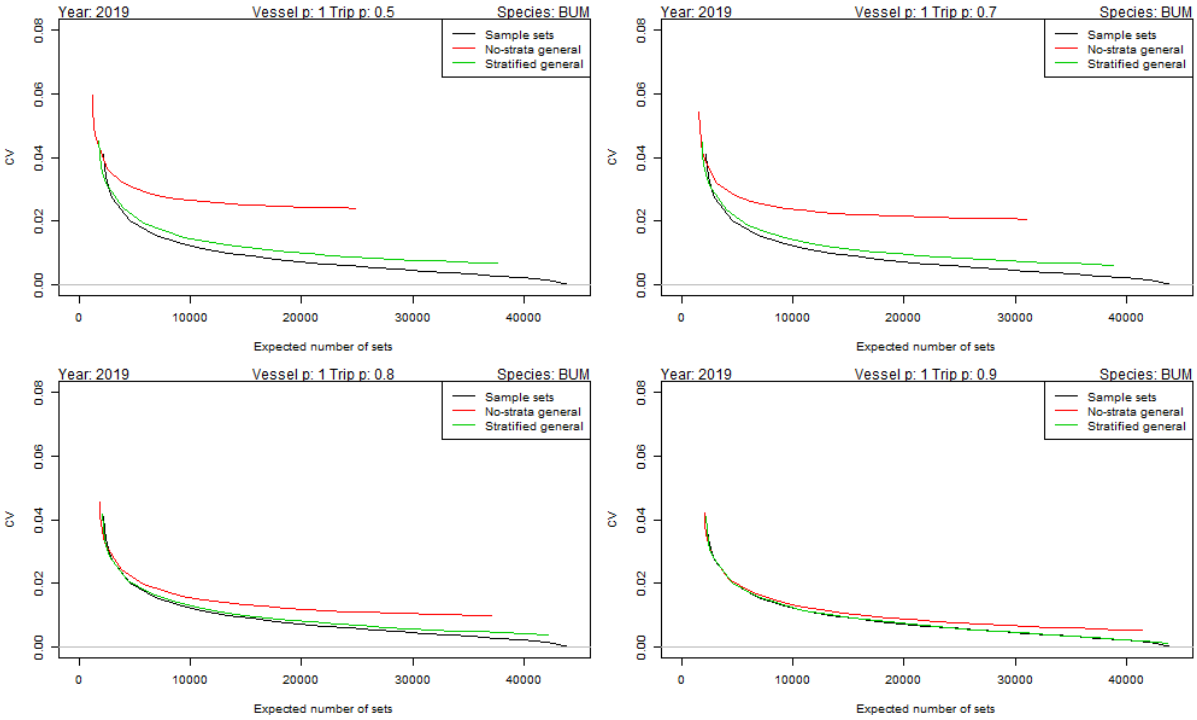

Figure 25: Longline logbook data: BUM, 10° S to 20° N, 2019: the CV of total catch vs the expected number of sets sampled for four different levels of trip selection (Trip p = 0.5, 0.7, 0.8, 0.9) and sampling with or without trip stratification. In each case all vessels were sampled (Vessel p = 1). The constant reference curve is shown in black for sampling a proportion of sets for all trips (Sample sets).

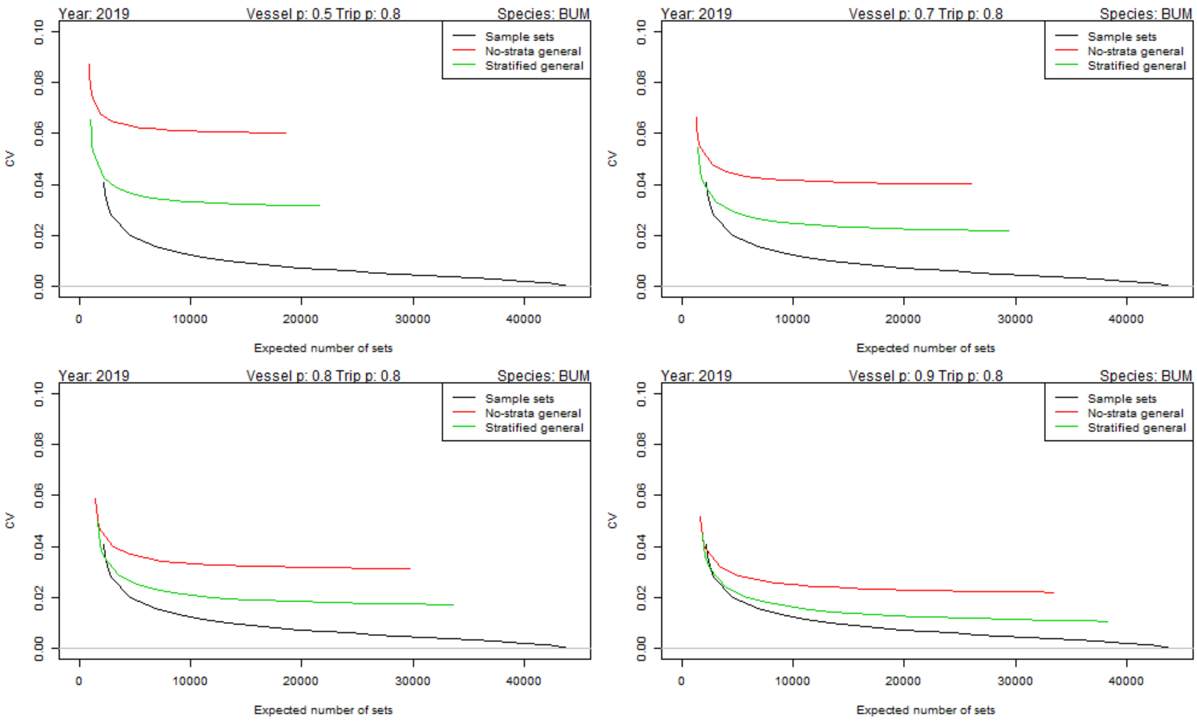

Figure 26: Longline logbook data: BUM, 10° S to 20° N, 2019: the CV of total catch vs the expected number of sets sampled for four different levels of vessel selection (Vessel p = 0.5, 0.7, 0.8, 0.9) and sampling with or without trip stratification. In each case 80% of trips were sampled (Trip p = 0.8) for each selected vessel. The constant reference curve is shown in black for sampling a proportion of sets for all trips (Sample sets).

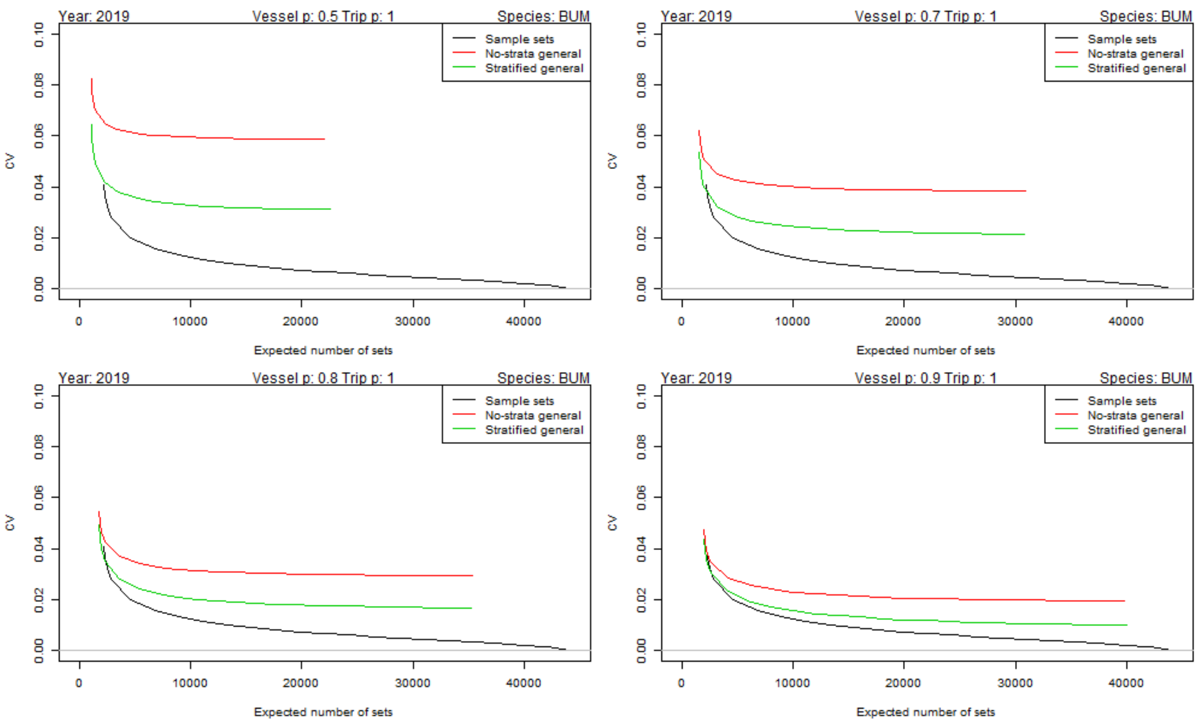

Figure 27: Longline logbook data: BUM, 10° S to 20° N, 2019: the CV of total catch vs the expected number of sets sampled for four different levels of vessel selection (Vessel p = 0.5, 0.7, 0.8, 0.9) and sampling with or without trip stratification. In each case all trips were sampled (Trip p = 1) for each selected vessel. The constant reference curve is shown in black for sampling a proportion of sets for all trips (Sample sets).

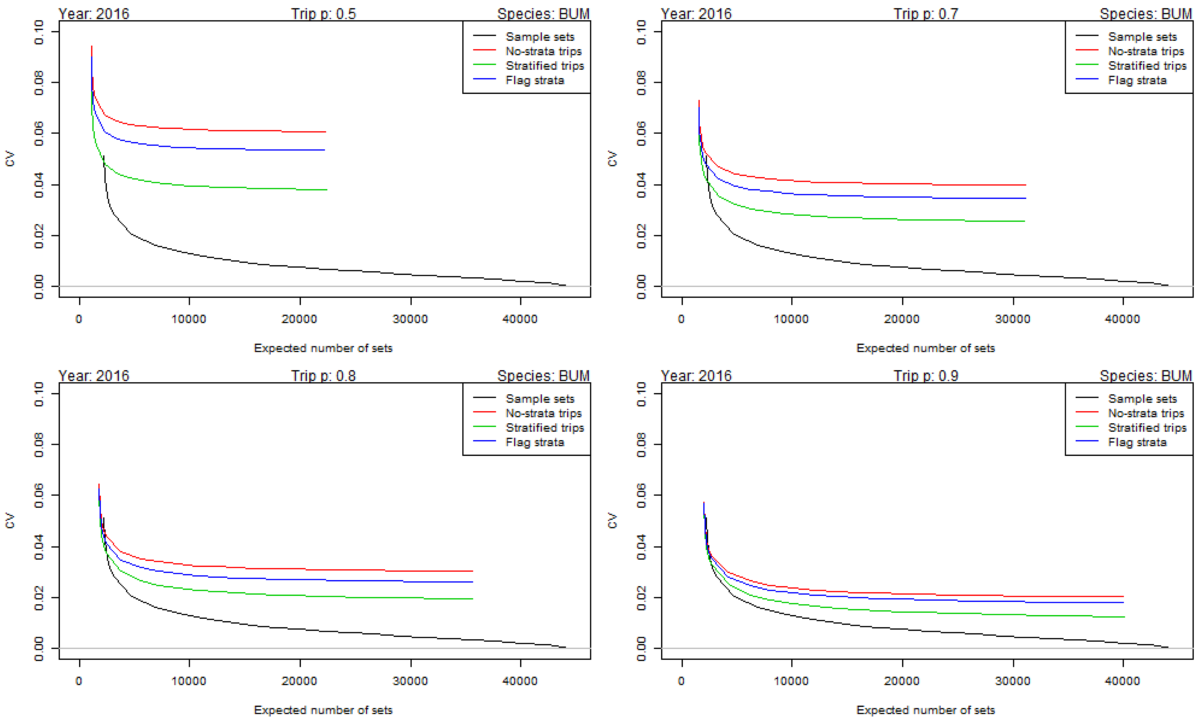

Figure 28: Longline logbook data: BUM, 10° S to 20° N, 2016: the CV of total catch vs the expected number of sets sampled for four different levels of trip selection (Trip p = 0.5, 0.7, 0.8, 0.9) and sampling with or without trip stratification, or with stratification by flag. In each case vessels were ignored. The constant reference curve is shown in black for sampling a proportion of sets for all trips (Sample sets).

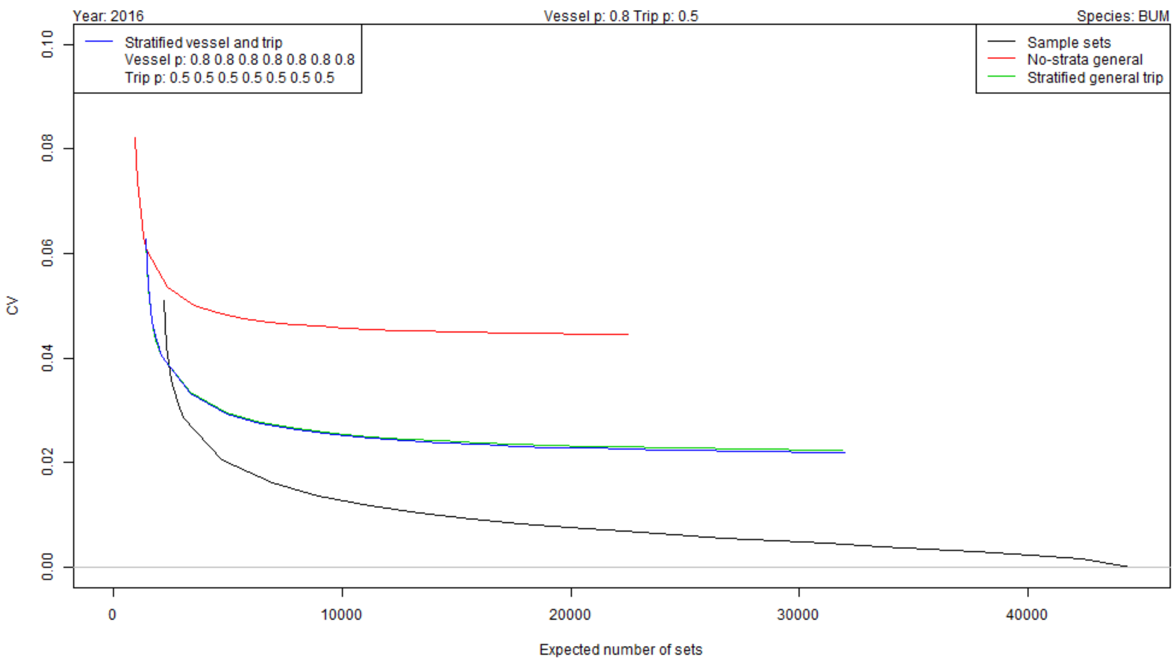

Figure 29: Longline logbook data: BUM, 10° S to 20° N, 2016: the CV of total catch vs the expected number of sets sampled for four different sampling scheme: sampling a constant proportion of sets for all trips (Sample sets); sampling 80% of vessels and 50% of the selected vessel’s trips with and without stratification: no stratification (No-strata general), by trip (Stratified general trip), by vessel and trip (Stratified vessel and trip).

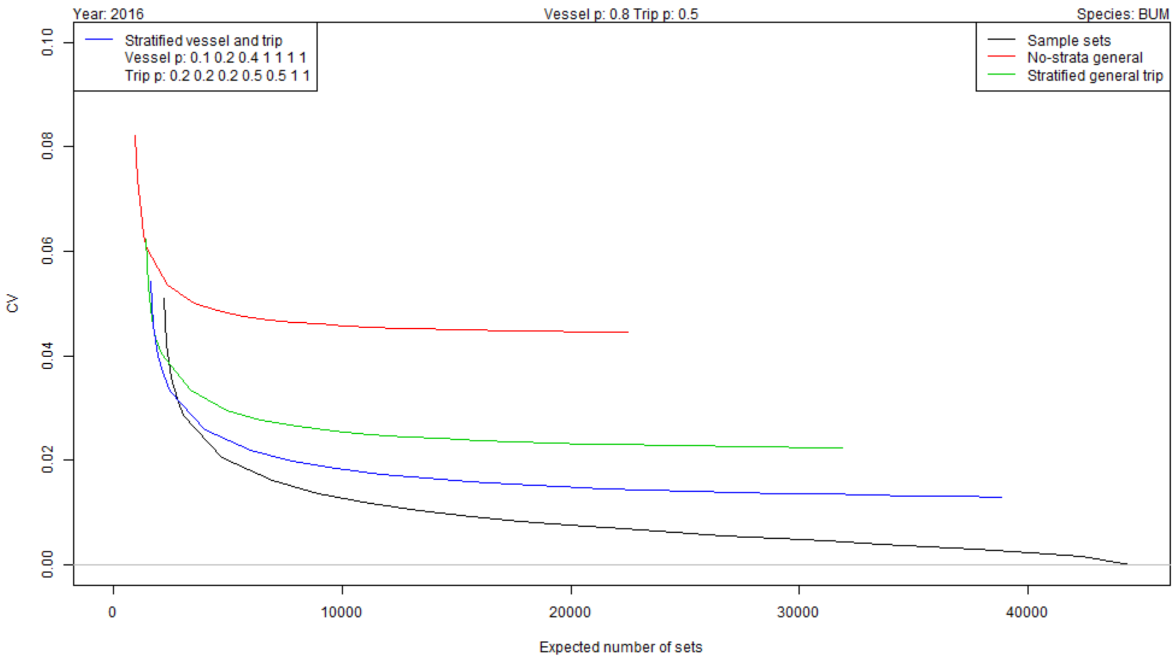

Figure 30: Longline logbook data: BUM, 10° S to 20° N, 2016: the CV of total catch vs the expected number of sets sampled for four different sampling scheme: sampling a constant proportion of sets for all trips (Sample sets); sampling 80% of vessels and 50% of the selected vessel's trips with and without stratification: no stratification (No-strata general), by trip (Stratified general trip); stratification by vessel and trip with different levels of vessel and trip proportion by strata (Stratified vessel and trip, vessel and trip proportions by strata as given).

Determination of appropriate sampling proportion

A key problem in determining appropriate sampling proportions for sets is that the precision achieved for a given sampling proportion varies greatly across species. Also, CVs are not a uniformly good measure of precision for estimators. When the mean of an estimator is very low then what appears to be a large CV is not actually a problem. For example, if the annual catch for a given species in a given area is 1 million fish, then a CV of 0.5 for an estimator of the annual catch would be completely inadequate; but, if the annual catch was 5 fish then a CV of 0.5 is adequate for most purposes.

The preliminary analyses and the investigation of alternative stratifications has shown that it is not generally possible to do any better than just sampling a given proportion of sets from each trip. The results below are for such estimators and the purpose of the plots is to decide which proportion of sets will deliver acceptable estimators for annual catches over large areas and more generally for catches within smaller areas or reduced time frames.

Longline

The first four plots below are for the 15 species that were chosen for longline, being 5 target species, which have catches adequately recorded in the logbooks, and 10 bycatch species for which observer catches were used to create operating models on the same scale as the logbooks. For the bycatch species, 200 operating models were generated and the median of the total catches and CVs across the operating models with total catch of at least 1 animal were used in the plots.

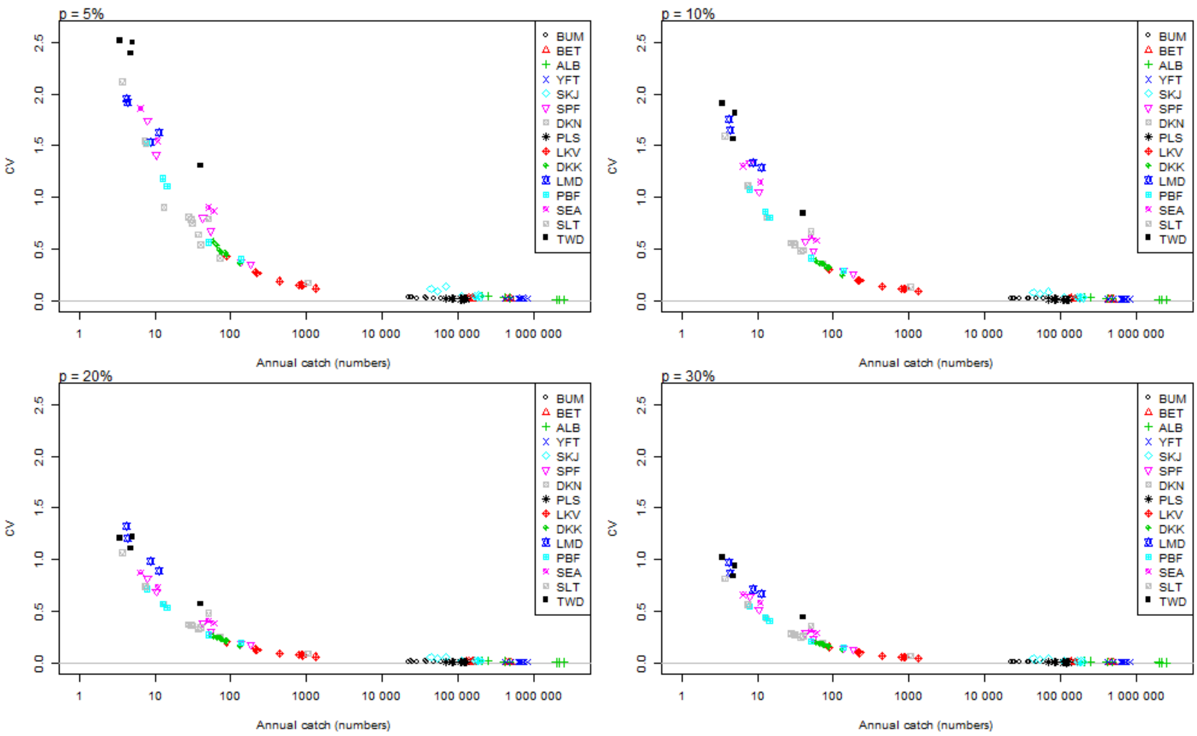

Figure 31: Longline species: CVs of annual catch estimators when 5, 10, 20, or 30% of sets are sampled from all trips. Points are plotted for the estimators of annual catches in 2016, 2017, 2018, and 2019 for the two main areas for twelve of the species. For TWD and SEA the area is south of 30° S and for LMD the area is north of 10° N.

The CVs for annual catches appear to be close to zero for the target species over the two main areas for sampling proportions from 5-30% (Figure 31). However, this is because of the scale of the y-axis and on closer examination, at p = 5% SKJ has two CVs that are above 0.1 (Figure 32). The CVs for bycatch species (including the three species outside the main areas) get well above 1 and up to about 2.5 for p = 5% (Figure 31). If there was a requirement across all 15 species to have a CV of less than 10% then the sampling proportion would need to be well above 30% (and indeed probably close to 100% for some species). However, this is where it is important to note that for small annual catches the use of a strict CV threshold is inappropriate.

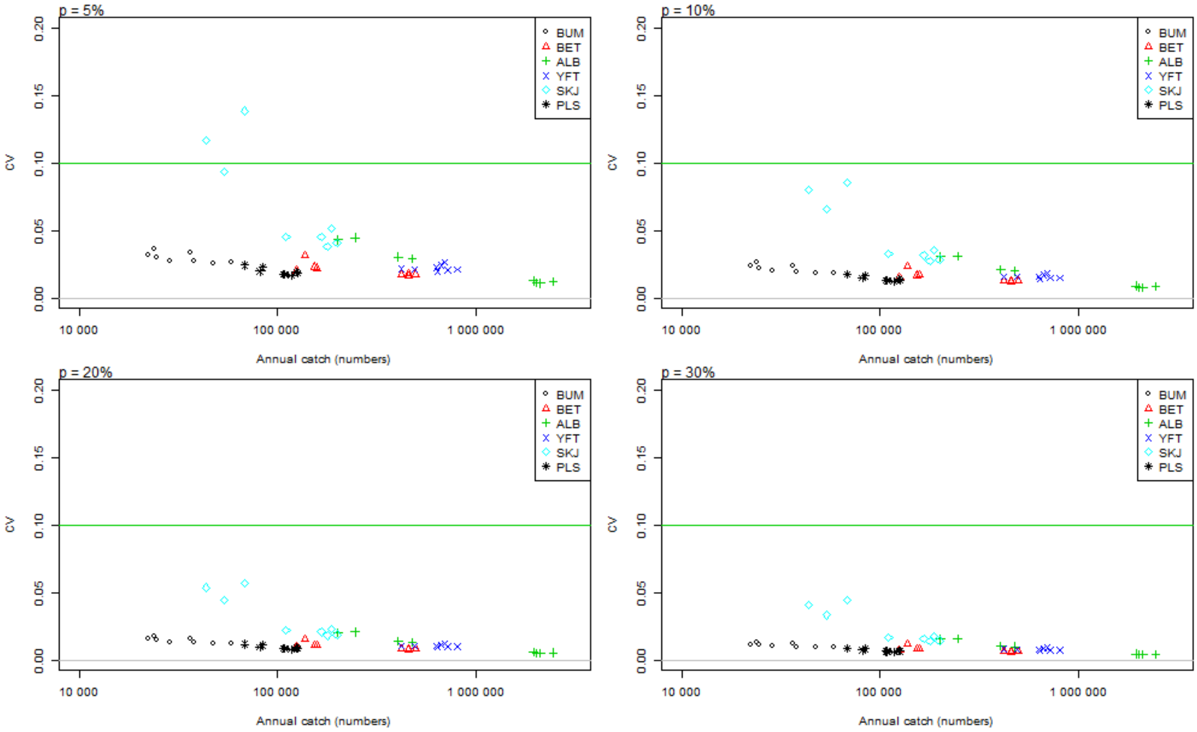

Figure 32: Longline species: CVs of annual catch estimators when 5, 10, 20, or 30% of sets are sampled from all trips. Points are plotted for the estimators of annual catches in 2016, 2017, 2018, and 2019 for the two main areas for species with an annual catch of more than 10 000 animals.

For most species with annual catches less than 10 000 animals, the ranges covering 95% of the distribution of the estimators of annual catch are not particularly wide even at p = 5% (Figure 33 but note that both axes have a log scale). But how “wide” is “acceptable”?

In Figure 34, an “acceptable” range is plotted for the width of 95% of an estimator’s distribution in terms of the ratio of the high point to the mean and the low point to the mean. The acceptable range was defined for mean catches from 1 to 10 000 animals using the table below:

Mean catch	Factor	Range
1	20	0.05-20
10	10	1-100
100	5	20-500
1000	2.5	400-2500
10 000	1.25	8000-12 500

For example, this says that for a catch of 10 animals that acceptable precision is achieved if at least 95% of the estimator’s distribution is within a factor of 10 of the mean (i.e., from 1-100 animals).

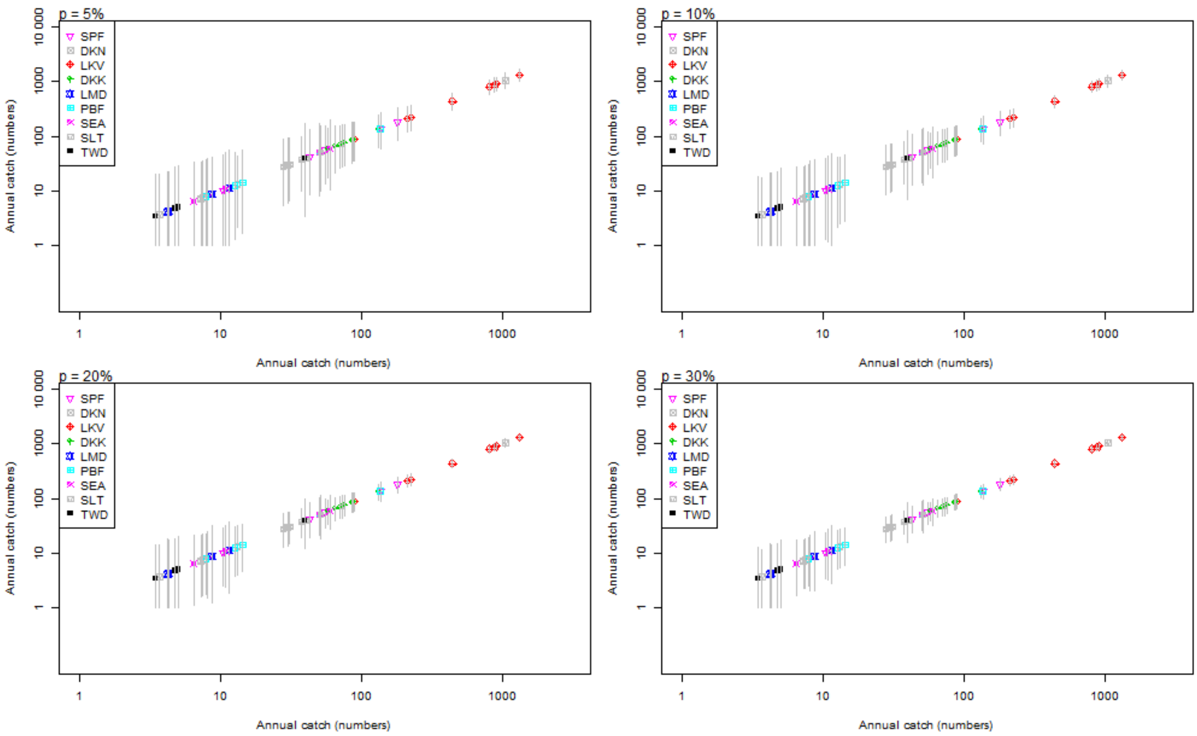

Figure 33: Longline species, annual catch less than 10 000 animals: the 95% two-sided range for the distribution of the estimators of annual catch when 5, 10, 20, or 30% of the sets from all trips are sampled. Years and areas as for Figure 31. A lognormal distribution was assumed. The lower 95% bound was set to 1 for intervals that extended below 1.

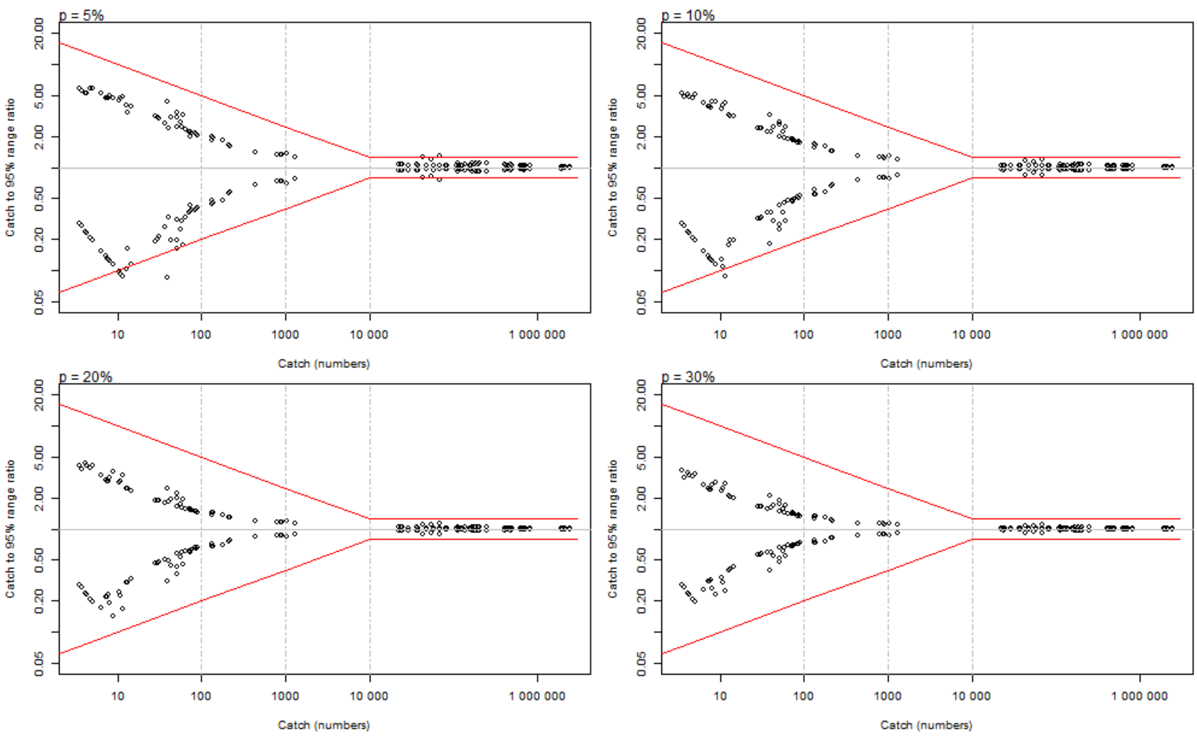

Figure 34: Longline species: the 95% two-sided range for the distribution of the estimators of annual catch when 5, 10, 20, or 30% of the sets from all trips are sampled. Years and areas as for Figure 31. The lower and upper points are plotted for each range as a ratio of the mean of the distribution. The red lines enclose the area of "acceptable" ratios (see the text for an explanation).

For the given definition of acceptability, then p = 5% is acceptable for almost all species (for the areas and years used) (Figure 34). Certainly, p = 10% is adequate for annual catches less than 10 000 animals (Figure 34) and for the target species (assuming that a CV of less than 10% is acceptable, see Figure 32).

The remaining four figures only consider the 5 target species (within the 15 selected) for longline. The annual catch for the 16 flag states, which had logbook catches in each of the four years, was considered for each of the two main areas and the four years. Also, the total catches were considered for different timeframes: the first 6 months, the first three months, and the first month (January). That is, for each of the 5 target species, CVs were calculated for the estimators of catch for each of the four years and the two areas, for annual catches, for catches in January, or the first 3 months, or the first 6 months, or the annual catch for any of the 16 flag states.

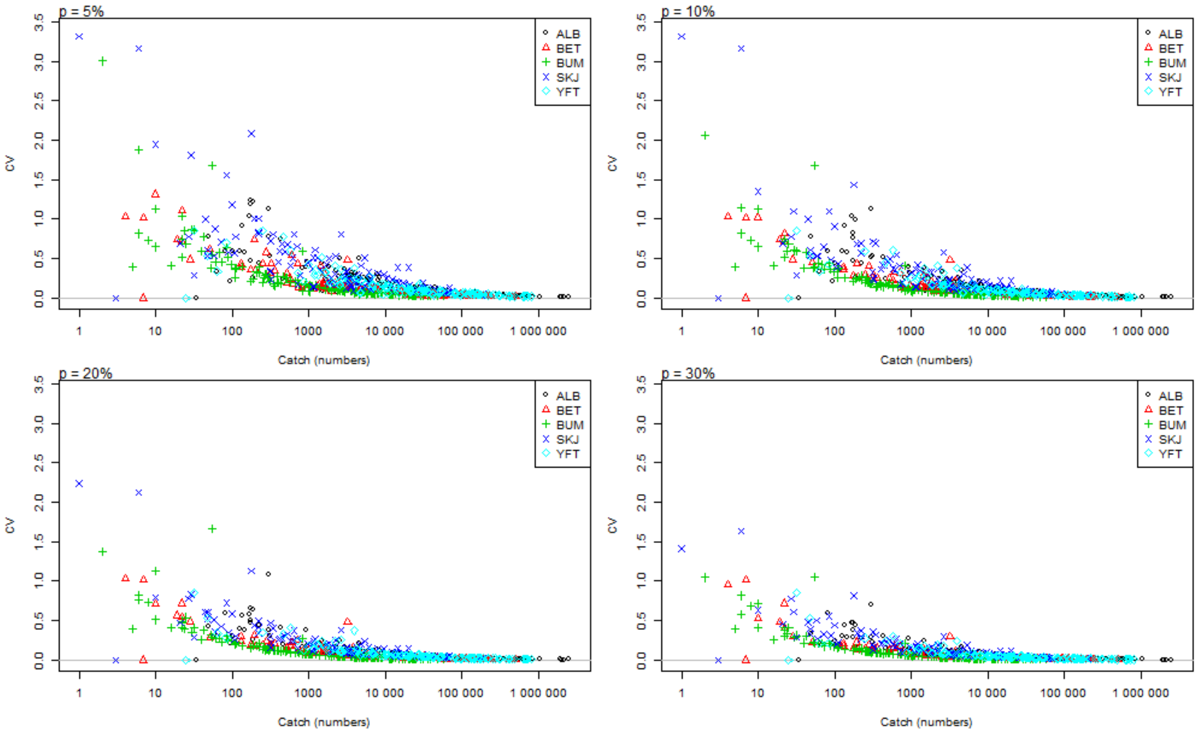

Figure 35: Target longline species: CVs of catch estimators when 5, 10, 20, or 30% of sets are sampled from all trips. Points are plotted for the estimators of catches in 2016, 2017, 2018, and 2019 for the two main areas and the five target species over different timeframes (annual, January, first 3 months, first 6 months) and annual catch for each of 16 flag states.

With the very broad collection of areas and timeframes the catch estimators have their CVs spread over a wide range of values (Figure 35). The curve is less well defined than it was for the annual catches of the 15 longline species (see Figure 31). For catches of more than 10 000 fish, it is necessary to have a sampling proportion of 20% to get the CVs mainly under 10% (Figure 36). Also, to get acceptable precision for the estimators of catches less than 10 000 fish, it is again necessary to have a sampling proportion of 20% (Figures 37 & 38). This is in contrast to the annual catches of the 15 longline species where a sampling proportion of 10% was adequate.

For the given definition of "acceptable" precision, it follows that acceptable precision for annual catches in the main areas can be achieved with a sampling proportion of 10%. However, if it is necessary to have adequate precision for the catches of species in smaller areas and/or less than annual timeframes, and/or for flag states, then a sampling proportion of 20% is required.

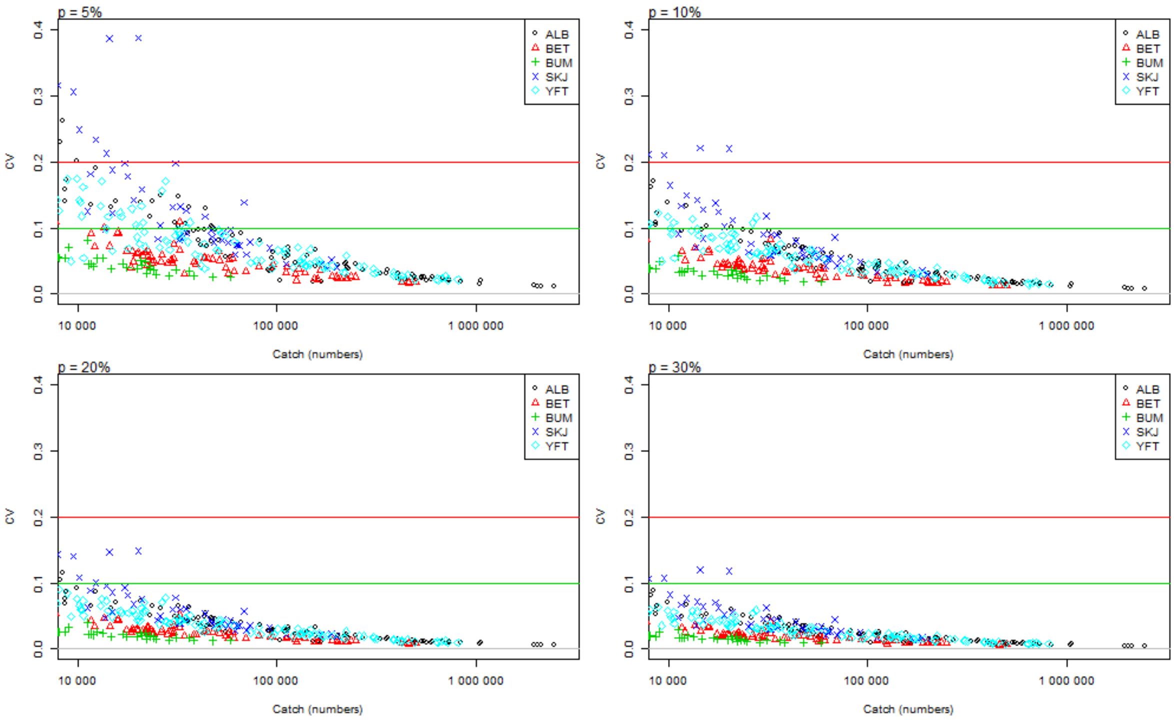

Figure 36: Target longline species: CVs of catch estimators when 5, 10, 20, or 30% of sets are sampled from all trips. Points are plotted for the estimators of catches in 2016, 2017, 2018, and 2019 for the two main areas and the five target species over different timeframes (annual, January, first 3 months, first 6 months) and annual catch for each of 16 flag states for catches of 10 000 animals or higher.

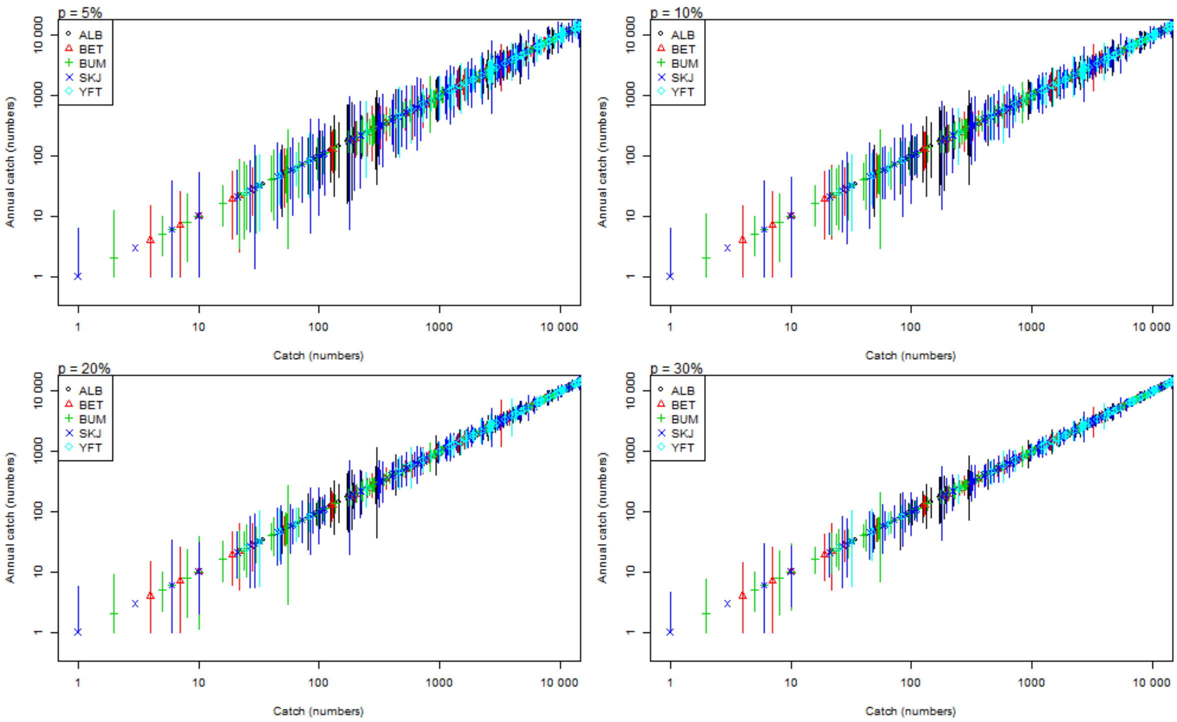

Figure 37: Target longline species, annual catch less than 10 000 animals: the 95% two-sided range for the distribution of the estimators of catch when 5, 10, 20, or 30% of the sets from all trips are sampled. Timeframes, areas, and flags as for Figure 5. A lognormal distribution was assumed. The lower 95% bound was set to 1 for intervals that extended below 1.

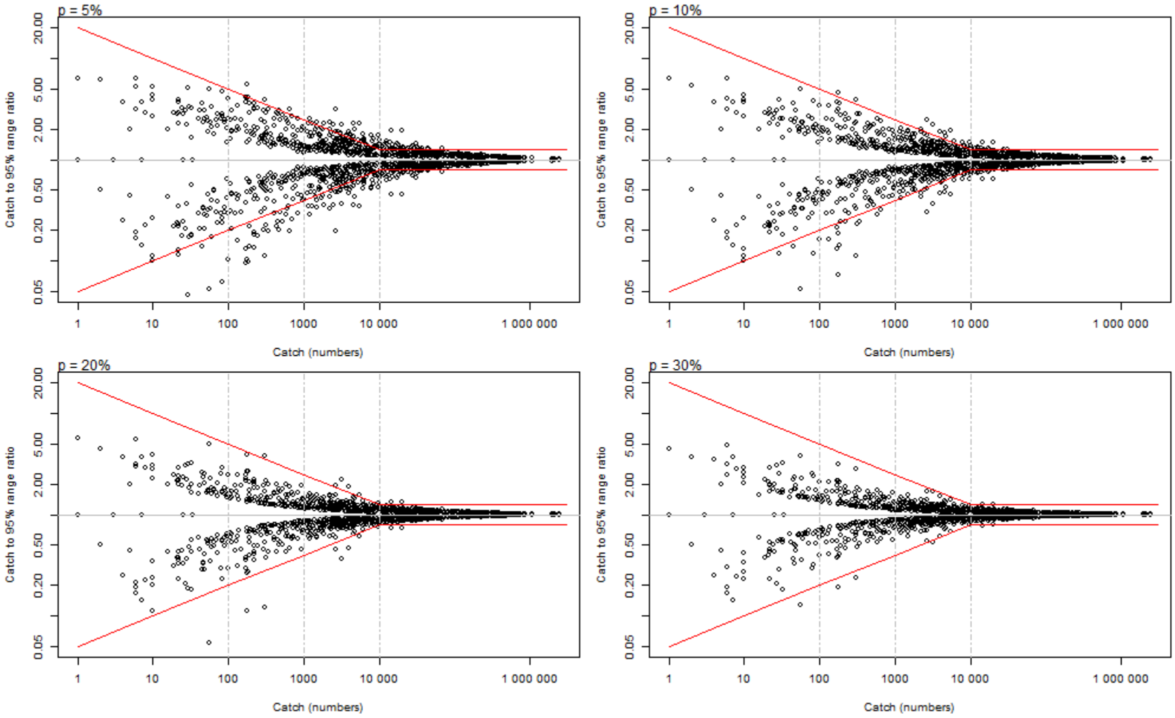

Figure 38: Target longline species: the 95% two-sided range for the distribution of the estimators of catch when 5, 10, 20, or 30% of the sets from all trips are sampled. Timeframes, areas, and flags as for Figure 35. The lower and upper points are plotted for each range as a ratio of the mean of the distribution. The red lines enclose the area of "acceptable" ratios (see the text for an explanation).

Purse seine

As for longline, similar plots were produced for the 15 species selected for purse seine, the years 2016-2019, and the two main areas (Figures 39-42). For target species in the northern area, sampling 10% of the sets for each vessel provides good precision for annual catches (Figure 40). However, for species in years and areas with lower annual catches, sampling 20% of sets is needed to generally get "adequate precision" (as defined for longline catches except that here the catches are in tonnes instead of numbers) (Figure 42). A similar conclusion was reached for longline.

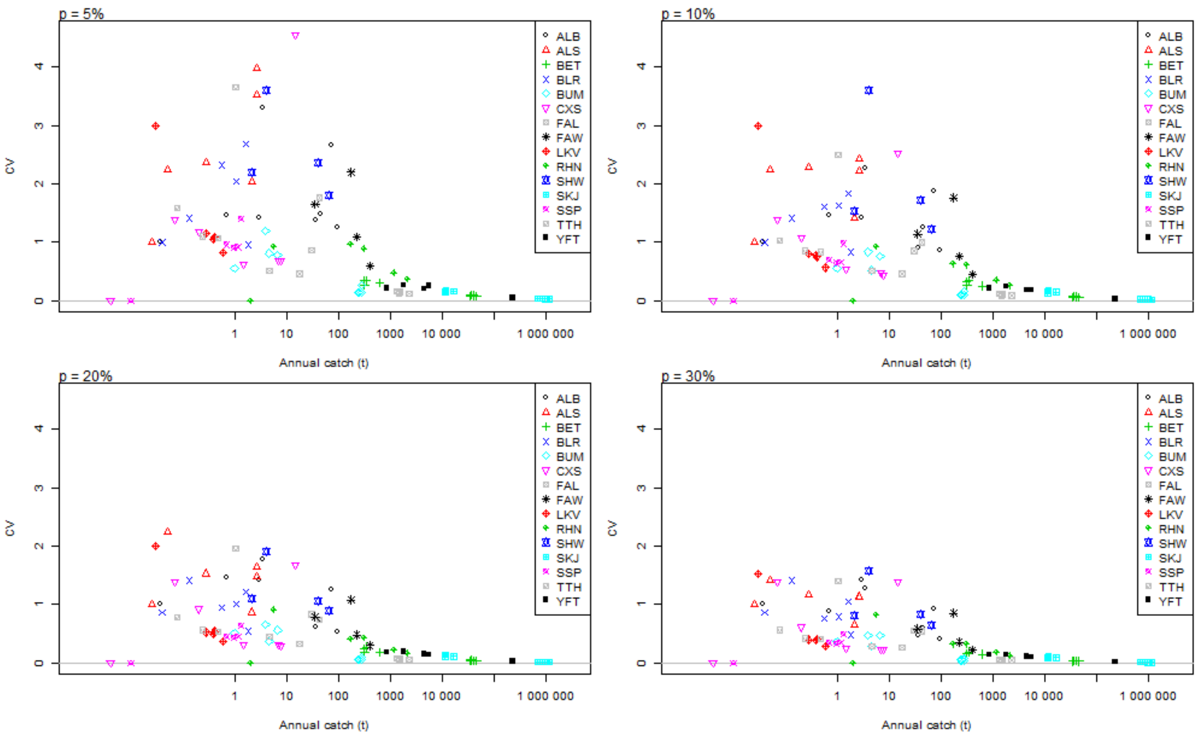

Figure 39: Purse seine species: CVs of annual catch estimators when 5, 10, 20, or 30% of sets are sampled from all trips. Points are plotted for the estimators of annual catches in 2016, 2017, 2018, and 2019 for the two main areas and fifteen species.

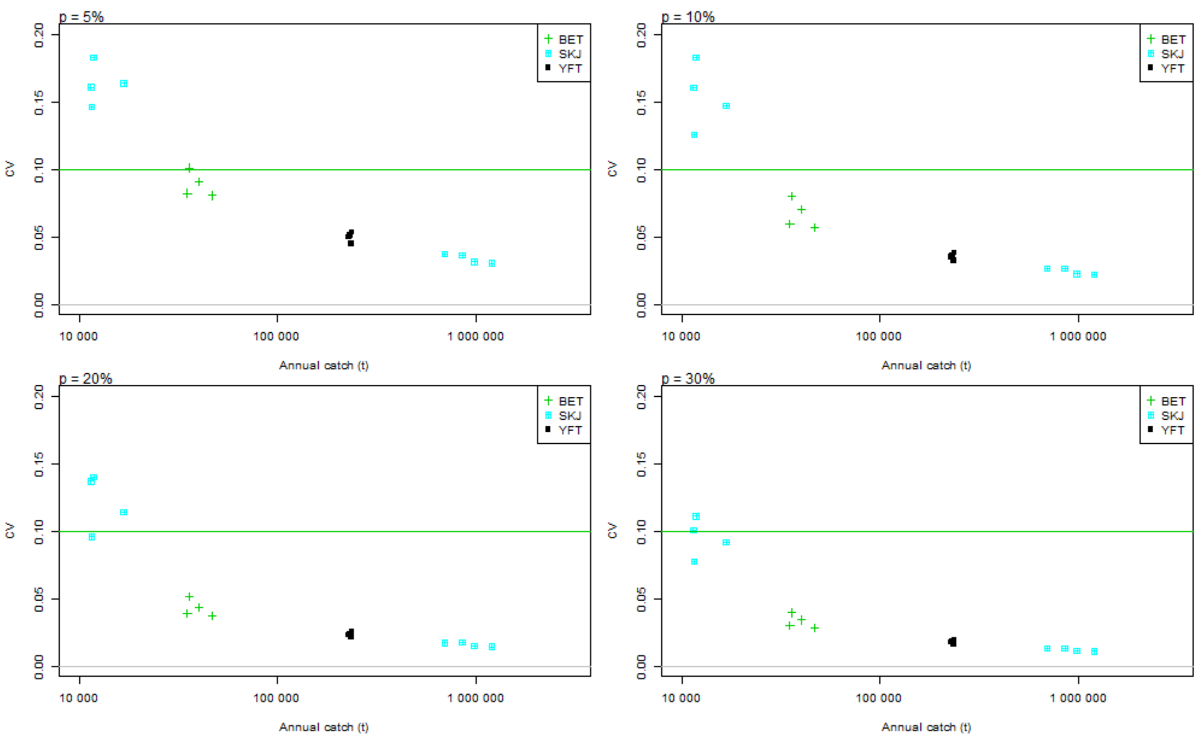

Figure 40: Purse seine species: CVs of annual catch estimators when 5, 10, 20, or 30% of sets are sampled from all trips. Points are plotted for the estimators of annual catches in 2016, 2017, 2018, and 2019 for the two main areas and species with an annual catch within an area of more than 10 000 t.

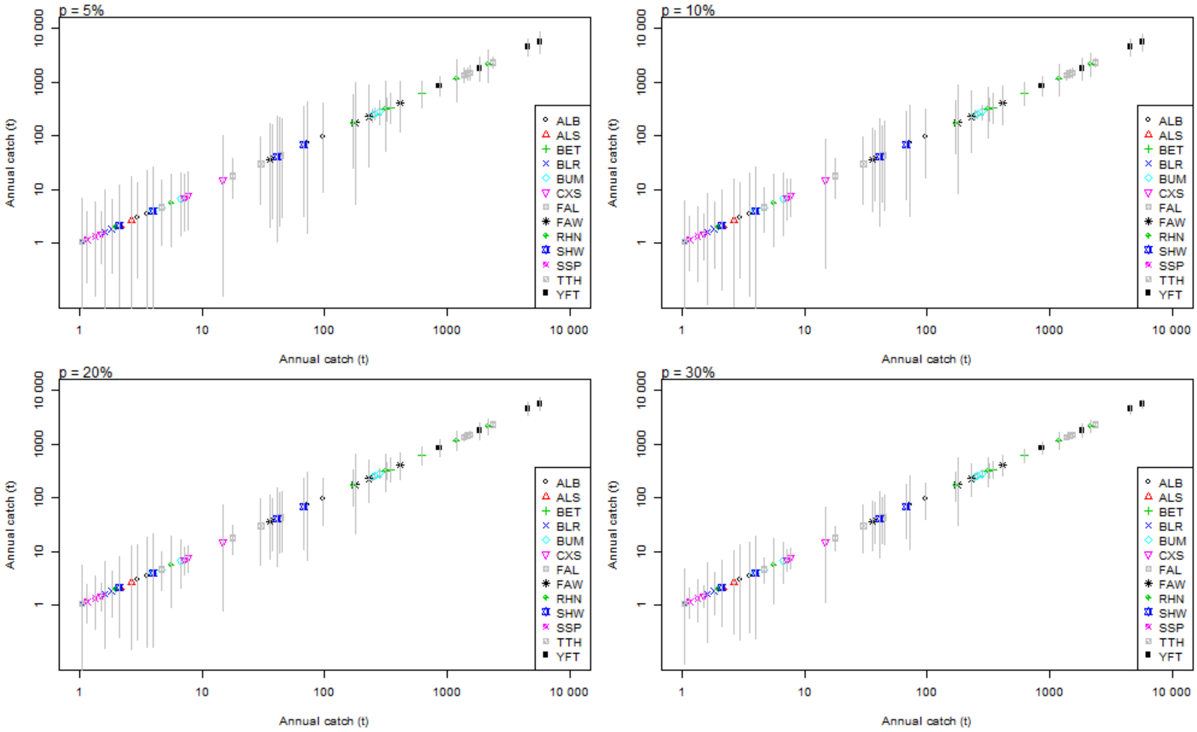

Figure 41: Purse seine species, annual catch less than 10 000 t: the 95% two-sided range for the distribution of the estimators of annual catch when 5, 10, 20, or 30% of the sets from all trips are sampled. Years and areas as for Figure 39. A lognormal distribution was assumed.

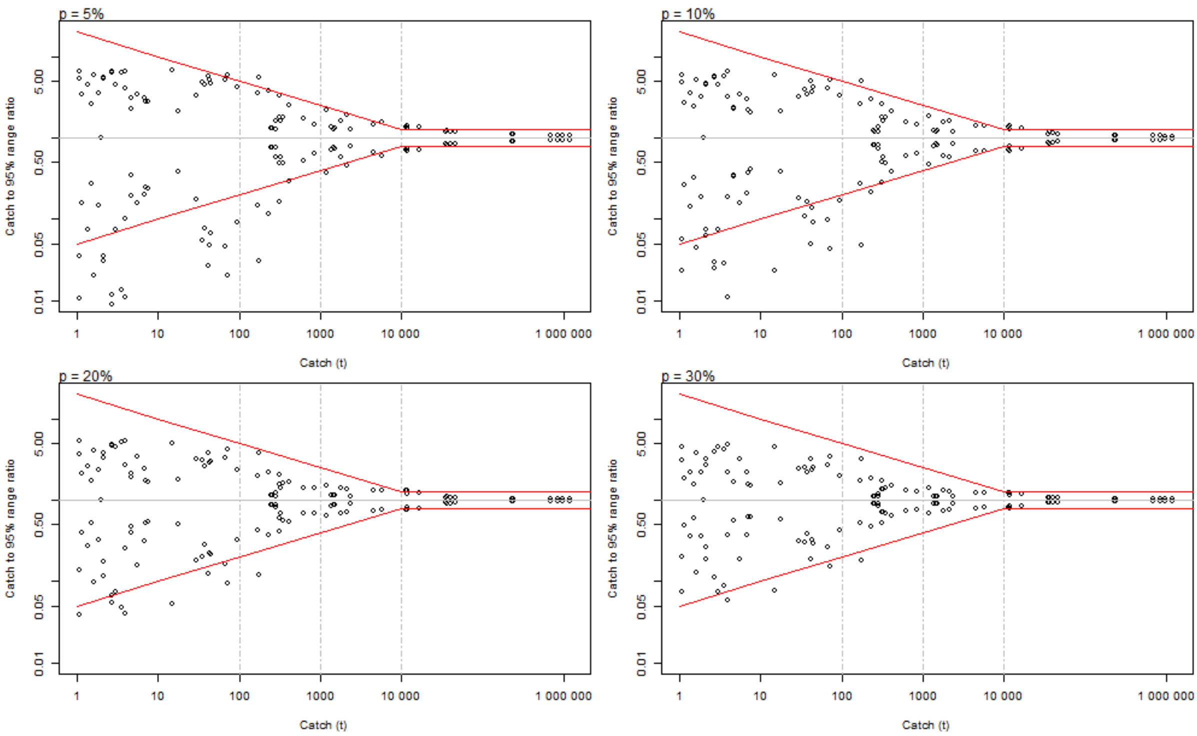

Figure 42: Purse seine species: the 95% two-sided range for the distribution of the estimators of annual catch when 5, 10, 20, or 30% of the sets from all trips are sampled. Years and areas as for Figure 39. The lower and upper points are plotted for each range as a ratio of the mean of the distribution. The red lines enclose the area of “acceptable” ratios.

Regional considerations

It seems that under most circumstances, in a regional setting, that a sampling proportion of 20% of sets should provide acceptable precision provided that the definition of "acceptable precision" used to derive the rule is appropriate. A regional manager may be faced with a particular set of circumstances which are not typical of what has already been considered or they may not accept the definition of "acceptable precision" used. This section aims at providing some guidance for different numbers of sets and the proportion of positive catches for a given species.

For bycatch species which are infrequently caught (5-20% positive catches) with the total number of sets 100 or greater, the distribution of the total catch estimator, for a given sampling proportion, can be approximated using a lognormal distribution. This can be done for any mean level of catch and, in particular, for a standardized mean catch of 1. The standardized distributions can be used to plot, for example, a range that covers 95% of the estimator's distribution for each sampling proportion from 0.01 to 1 for different numbers of sets (Figure 43). The point of a such a plot is that it can be used to answer questions like, "what sampling proportion is needed to be 95% confident that the estimate of the catch of a species will be within a factor of 5 of the true value"?

For example, consider a situation where a manager is expecting there to be about 200 sets (during some time interval and within some area) and they are interested in a species which is likely caught on about 5% of the sets. If they want to be 95% confident that the estimate lies within a factor of 3 then a sampling proportion of 20% is adequate (Figure 43, top right). They can put that range into absolute numbers by specifying an expected catch. For example, if 20 animals are caught, then a factor of 3 gives a range of about 6-60 animals. They can ask themselves if that sort of range is adequate for their purposes. If they wanted to be 99% confident that the estimate lies within a factor of 3 then they would need to move to a higher sampling proportion of about 30% (Figure 44, top right).

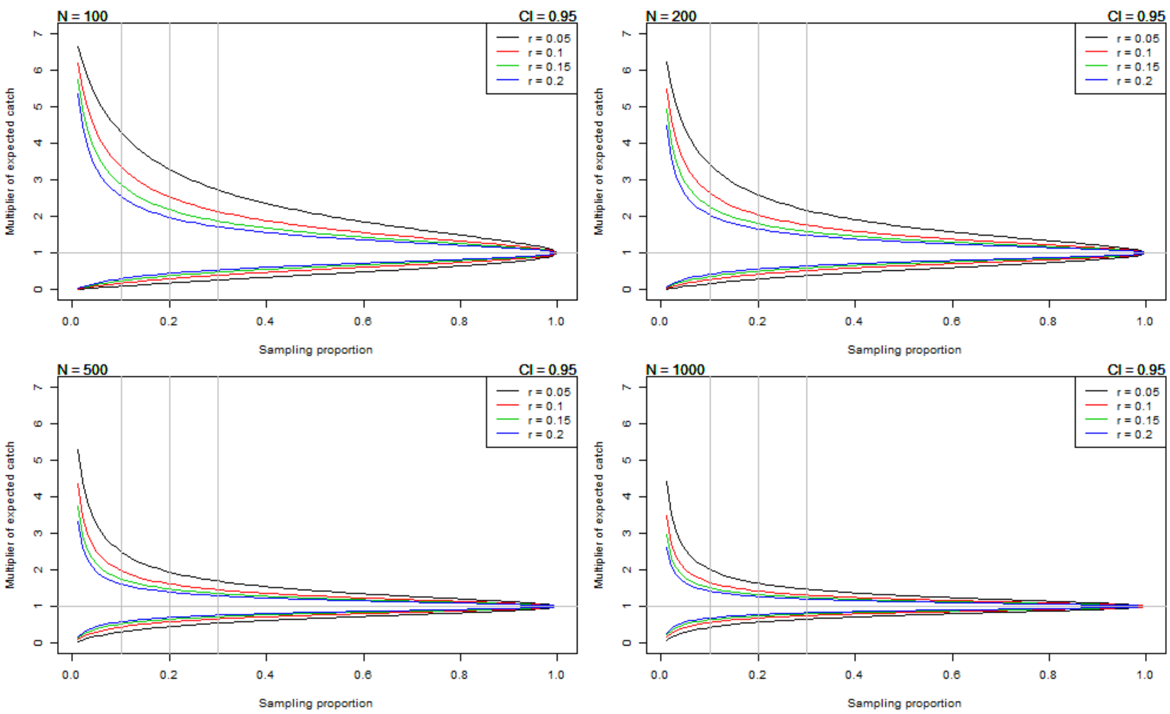

Figure 43: The multipliers of expected catch for 95% of the range of the total catch estimator's distribution. Values are shown for 100, 200, 500, or 1000 sets (*N*), proportion of positive catches from 5% to 20% (*r*), and sampling proportions from 1% to 100%. Vertical lines are plotted at 10%, 20%, and 30%. The horizontal line is at 1.

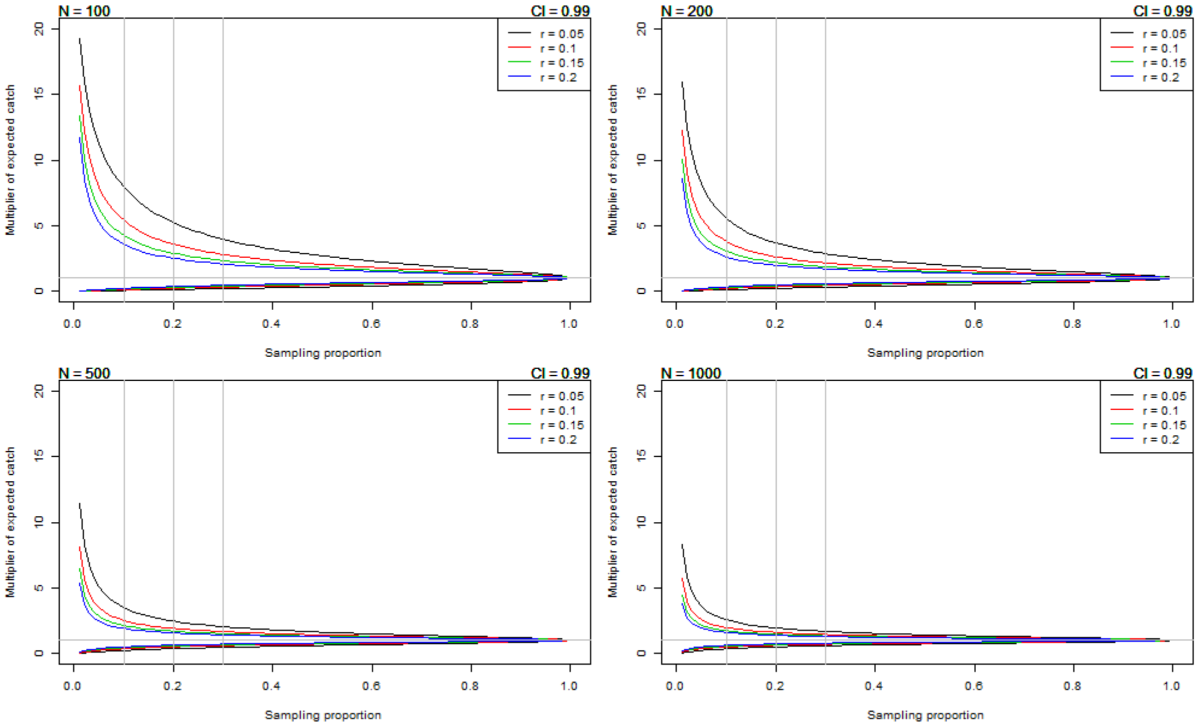

Figure 44: The multipliers of expected catch for 99% of the range of the total catch estimator's distribution. Values are shown for 100, 200, 500, or 1000 sets (*N*), proportion of positive catches from 5% to 20% (*r*), and sampling proportions from 1% to 100%. Vertical lines are plotted at 10%, 20%, and 30%. The horizontal line is at 1.

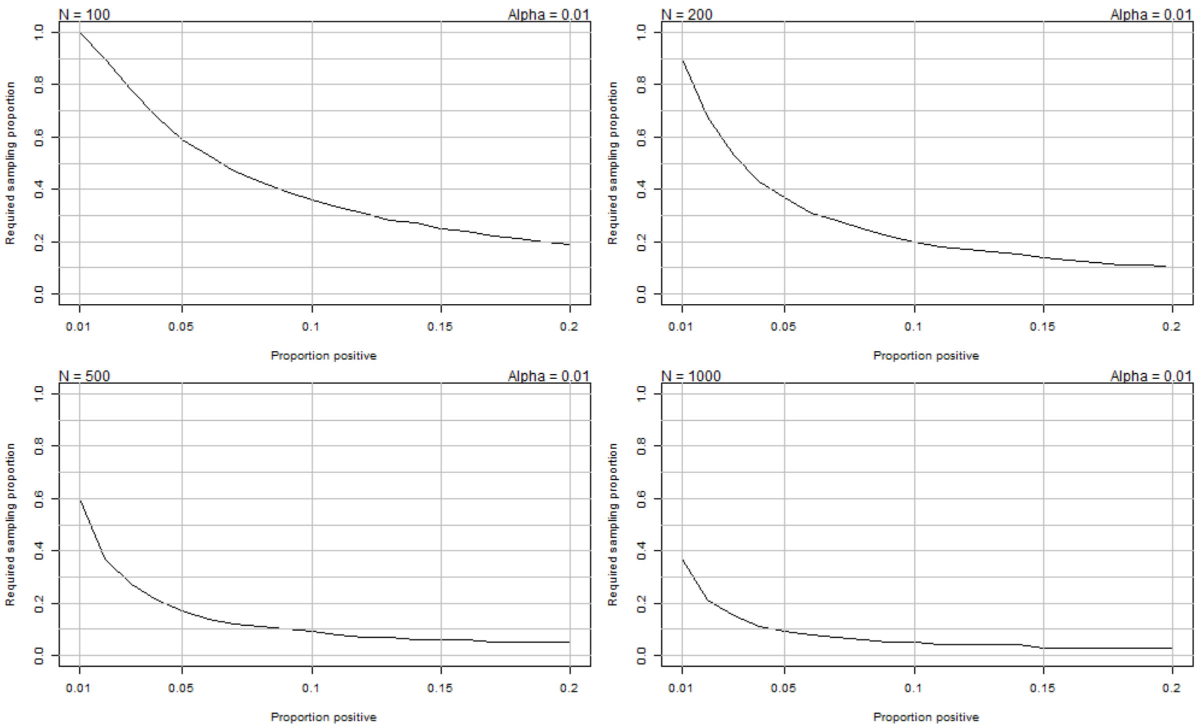

Figure 45: The sampling proportion required to reduce the probability of falsely observing a zero catch to 1% or lower (alpha). Values are shown for 100, 200, 500, or 1000 sets (*N*) and a proportion of positive catches from 1% to 20%.

For a species of special interest, in some area over some timeframe, a manager may need to know whether any of the species was caught or not. In this case, it is not a question of precision as such but a need to have a low probability of falsely reporting a zero catch. When a series of catches is viewed as a sequence of zero and non-zero catches and it is sampled a random without replacement then the number of non-zero catches obtained follows a hypergeometric distribution. The probability of a zero observed catch can be calculated using this distribution. For a given number of sets, a given probability threshold, and a given proportion of sets with a positive catch, the sampling proportion required to get the probability of a zero estimate below the threshold can be calculated (Figures 45-48).

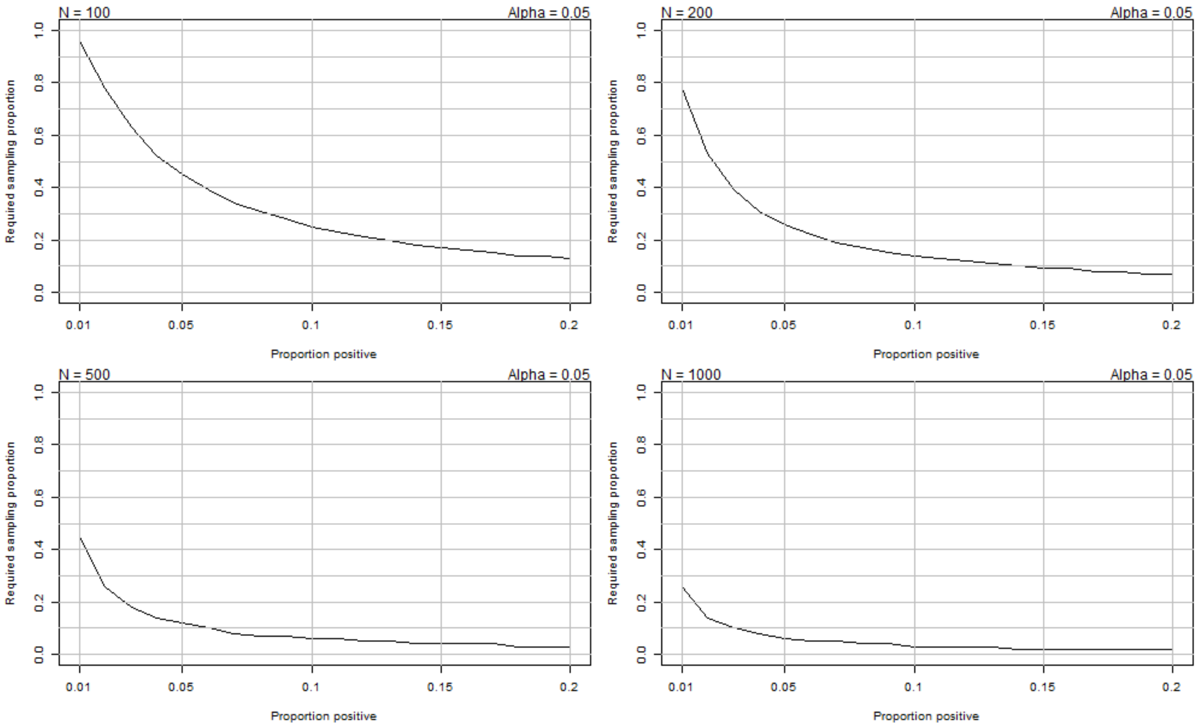

Figure 46: The sampling proportion required to reduce the probability of falsely observing a zero catch to 5% or lower (alpha). Values are shown for 100, 200, 500, or 1000 sets (*N*) and a proportion of positive catches from 1% to 20%.

For example, consider the situation where a manager is expecting about 500 sets (in a given area and timeframe) and wants to be 95% confident about whether a species has been caught or not. If the species is generally caught in about 1% of the sets, then a sampling proportion of about 45% is required (Figure 46, bottom left). If the species is generally caught in about 5% of the sets, then a sampling proportion of 15% would be adequate (Figure 46, bottom left).

For a species which is rarely caught it may be necessary to sample all the sets to know whether the species was caught or not. For example, if there are only 100 sets and the species is generally caught on less than 1% of the sets then it would be necessary to sample all the sets to be 99% confident of whether the species was caught or not (Figure 45, top left). Even when there are a larger number of sets, there will still potentially be species that are so rarely caught that a large proportion of the sets will need to be sampled to be confident of whether the species was caught or not (e.g., Figure 48).

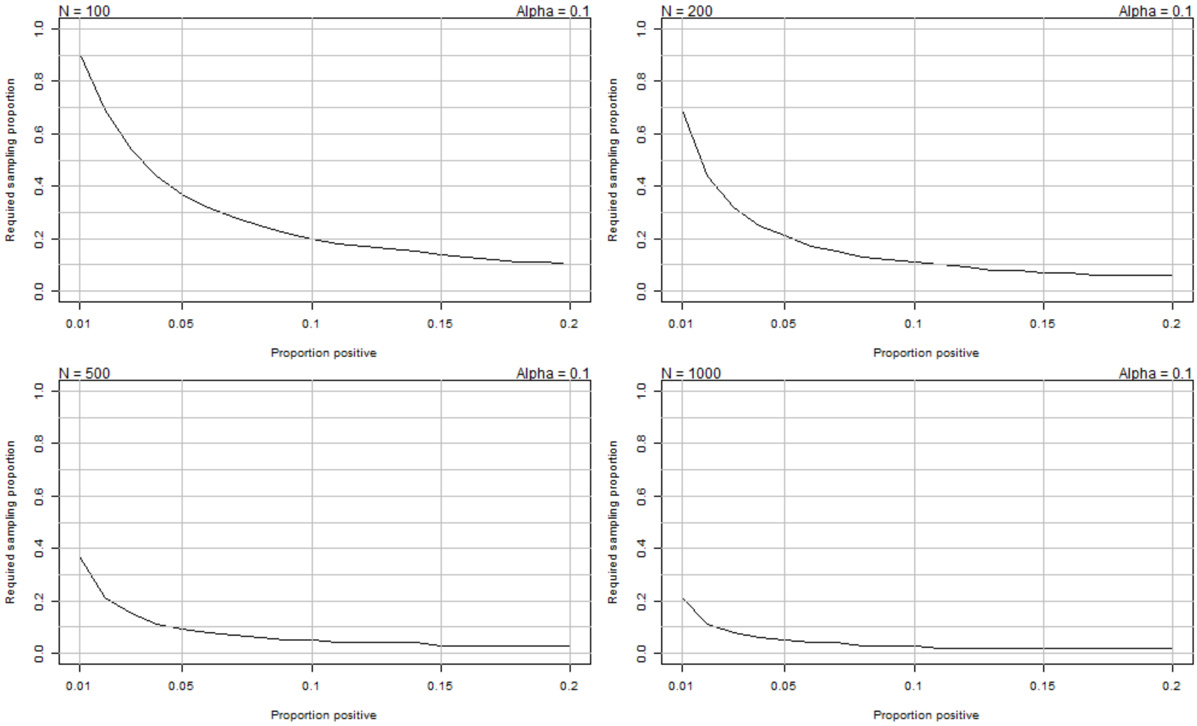

Figure 47: The sampling proportion required to reduce the probability of falsely observing a zero catch to 10% or lower (alpha). Values are shown for 100, 200, 500, or 1000 sets (*N*) and a proportion of positive catches from 1% to 20%.

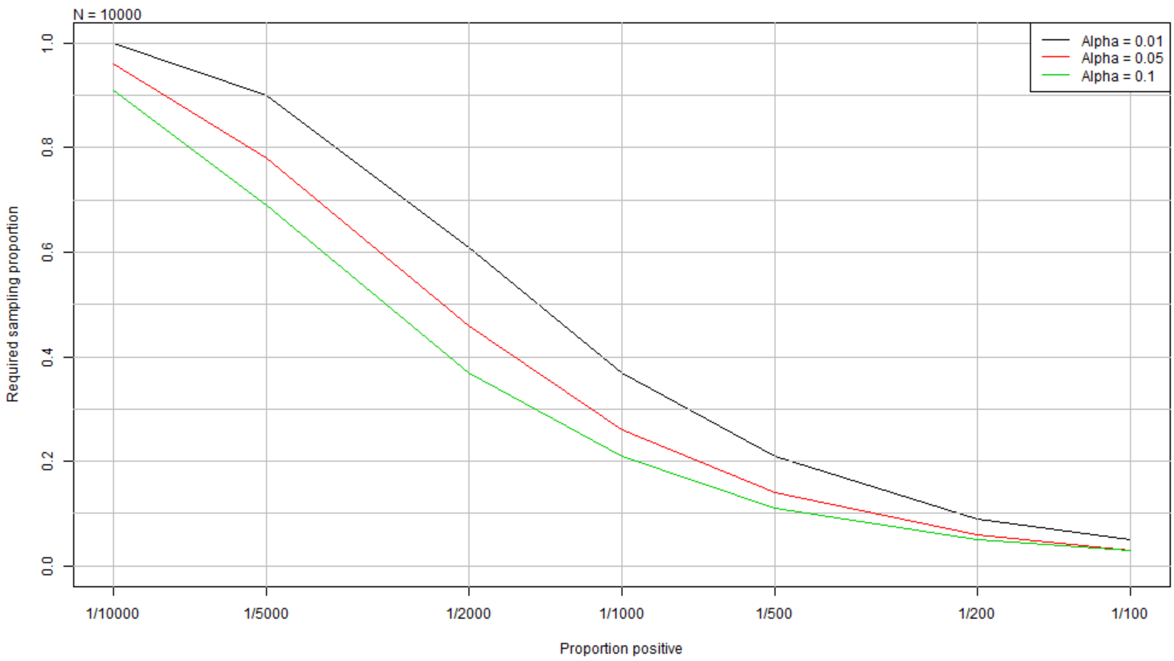

Figure 48: The sampling proportion required to reduce the probability of falsely observing a zero catch to 1%, 5%, or 10% when there are 10 000 sets. Values are shown for proportion positive ranging from 1 in 10 000 to 1 in 100 (on a log scale).

Sampling of hooks between floats

The field "hooks between floats" (hkbf) is important for two reasons. It is multiplied by the number of baskets to give the total number of hooks set and it is also used with the hook number, that an animal was caught on, to determine the depth of the hook that caught the animal. It is not important in terms of an average or a total for the fleet. It is mainly of interest for each particular set and only of minor interest in terms of the range of values that it may take. Therefore, the proportion of sets that are sampled by an EM analyst is not particularly relevant to the "precision" achieved for hkbf.

However, there are two questions with regard to the approach taken by EM analysts, that are relevant for hkbf. How should hkbf be determined for a set that is sampled and does it have to be separately determined for each set that is sampled in a trip? The current recommendation for EM analysts, for each set that they analyse, is to count the hooks in the first three baskets, the middle three baskets, and the last three baskets, and take an average of the nine numbers to provide a single hkbf (SPC et al. 2020). This is similar to the advice for observers although they are requested to give the most common value for hkbf if this changes during a set (SPC 2021).

The range of values recorded for hkbf for the observer data (2016-2019 inclusive) varies between 1 and 55 (Figure 49). The extreme values were checked to see if they were in error, but they appear to be genuine.

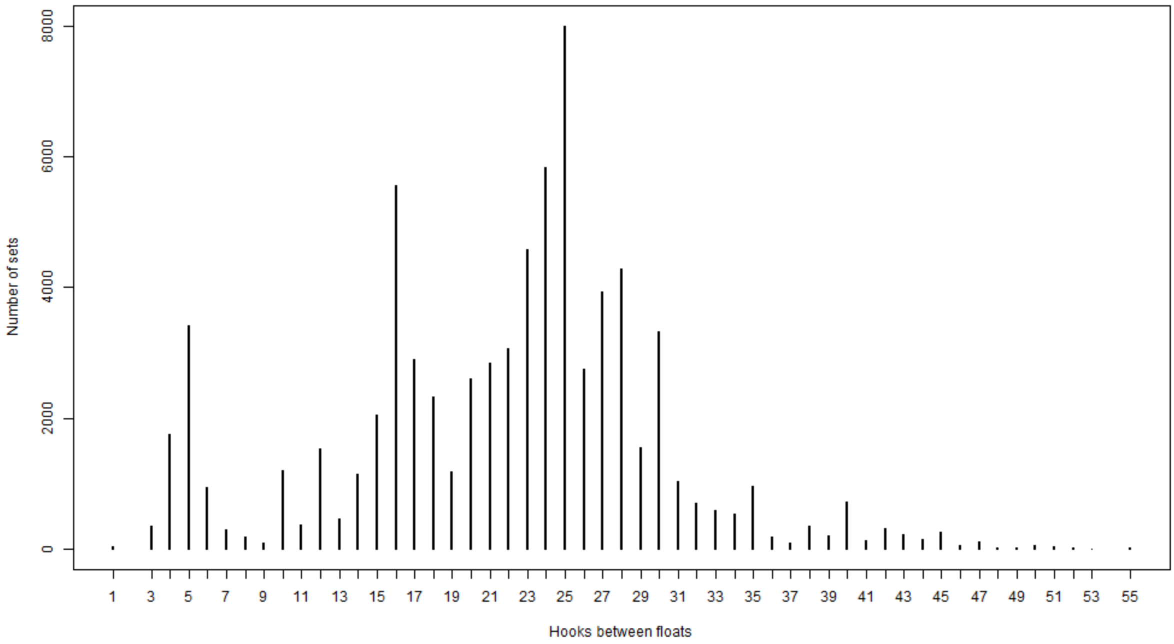

Figure 49: The frequency of recorded values for hkbf on longline sets in the 2016-2019 observer data.

One would expect that during a trip that a vessel may well set up the gear differently on different sets (to target different species by putting hooks at different depths through different hkbf). A quick look at the observer data confirms that this happens quite frequently: 71% of trips have the same hkbf for each set on a trip while 29% have at least two different values recorded. On some trips, up to 8 different values were recorded for hkbf (Figure 50). It seems clear that an EM analyst must count hooks between floats (in some fashion) for each set that they analyse within a trip.

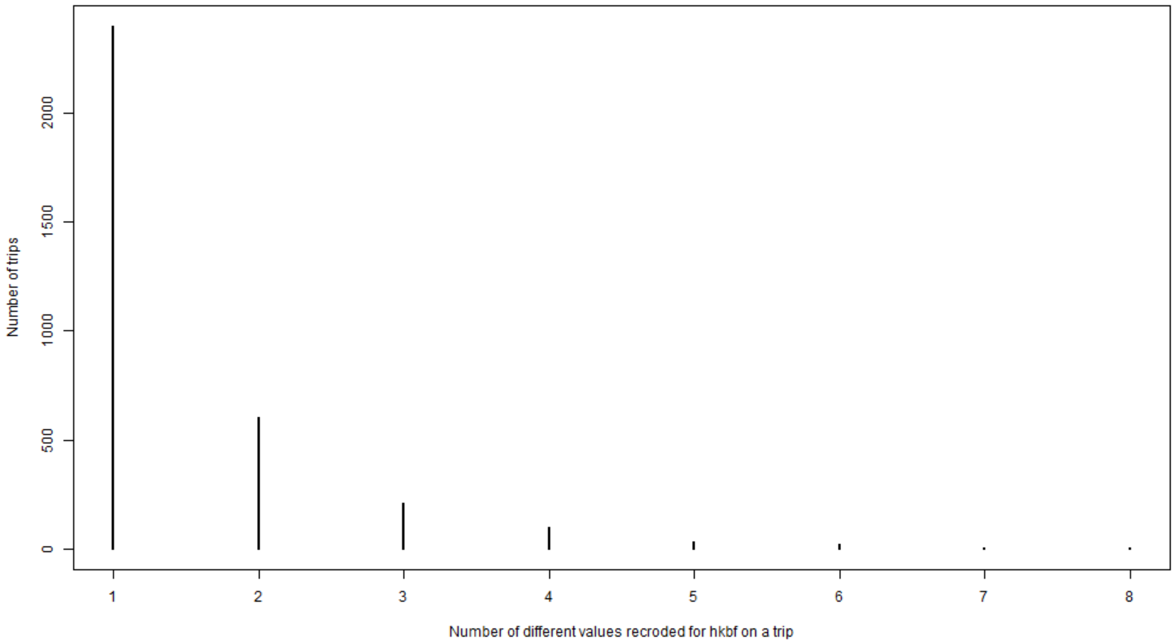

Figure 50: The frequency of longline trips on which different values were recorded for hkbf in the 2016-2019 observer data.

One would also think that hkbf could change during a set. The observer manual (SPC 2021) notes that this is the case and recommends that the observer look out for such changes as they affect the calculation of the total number of hooks set and also what value should be recorded for hkbf. The manual requests that the observer record the most frequent hkbf for these types of sets and note the details in a comment field. For about 20% of the sets in the observer data, the number of hooks sets is slighty different to the number of baskets set multiplied by hkbf (Figure 51). This is what would be expected if about 20% of the sets had a change in hkbf during the set.

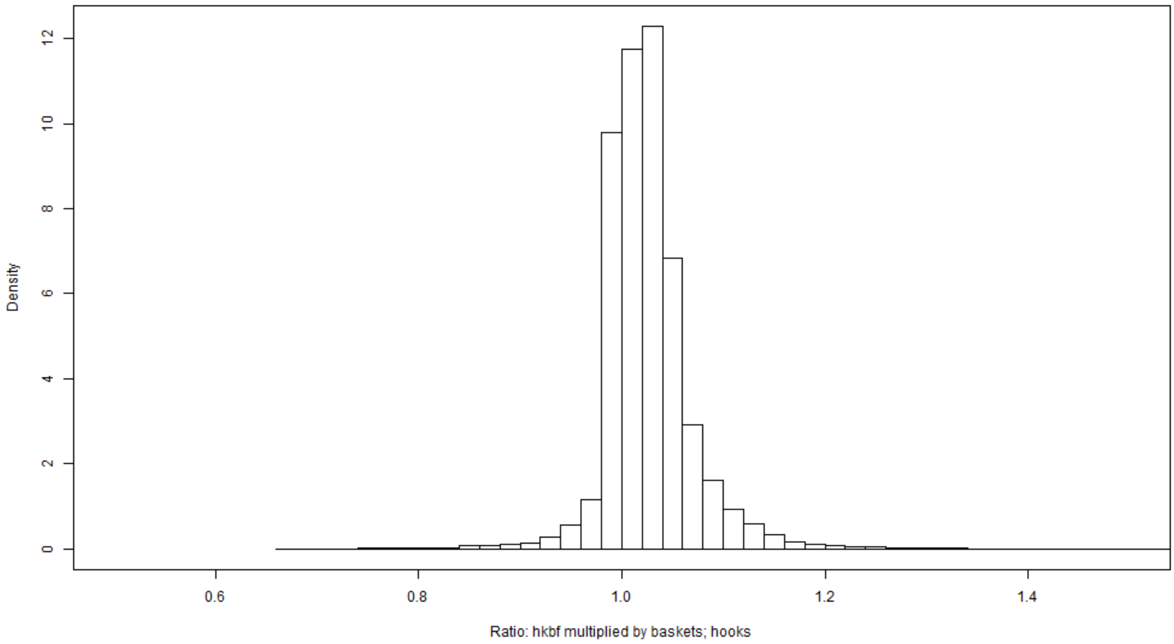

Figure 51: A histogram of the ratio of hkbf multiplied by the number of baskets, and the number of hooks set, for the longline observer data 2016-2019. A value different from 1 shows that the number of hooks set was not calculated by multiplying hkbf by the number of baskets set (implying that there was a change in hkbf during the set).

How then, should an EM analyst best go about counting hooks between floats on a given set? Just using the average of the first three, the middle three, and the last three baskets, seems inadequate. They could adopt the same approach as requested of the observers but this would mean that the detailed information would be "hidden" within a comment field. It seems more straightforward to allow an EM analyst to record more than one field for hkbf and the associated number of baskets. They have to look out for changes in the hkbf during a set in any case. They need the number of baskets associated with each hkbf to calculate the total number of hooks sets – since they need to note these values there is no reason why they shouldn't be recorded in the database. I would imagine that changes in hkbf occur within substantial blocks of baskets and that there wouldn't be more than two or three such blocks in any single set (perhaps there is some data on this in the observer comment fields).

This is not to say that an EM analyst needs to count the hkbf for every basket on each set that they analyse. They will need to do it for perhaps two or three baskets at the start of the set and then "look out" for any significant change in hkbf across the set. I can see that one way of "looking out" for a change in hkbf is to do the count for the first three baskets, the middle three, and the last three, but this hardly seems adequate. The key point is to identify each block of baskets that have a different number of hkbf. If this is not done correctly then the number of hooks set and the depth data for some captured animals could be quite inaccurate.

Discussion and recommendations

The current recommendation for EM analysts is to view a randomly selected proportion of sets for each trip and to target 20% of sets for vessels fishing in national waters and not landing into domestic ports with a minimum of 10% of sets viewed for domestic vessels (FFA 2020). This recommendation is based on the work of Lawson (2003, 2004) and Peatman & Nicol (2020). Lawson noted that the CV of an estimator of nominal CPUE decreased slowly as a function of the proportion of sets sampled from about 20%. Peatman & Nicol (2020) showed that sampling a proportion of sets from each trip was superior to sampling all sets from a proportion of trips.

In this report, a general sampling scheme involving a random selection of vessels, trips, and sets was investigated. It was found, when estimating total catch for a given species, that the simple approach of sampling a proportion of sets from every trip was almost always superior to any alternative approach. Even complex stratifications using information that would not generally be available at the time of stratification did not outperform the simple approach. The reasons are clear, as between-vessel and between-trip variation in catch can be huge. When all trips are selected between-trip and between-vessel variation does not impact on the estimator's precision.

Total catch was the main focus of the study as nominal CPUE will have very similar precision. This is because the CV for total hooks is small compared to the CV for total catch (when all trips are sampled). Intuitively this is clear as for a given trip there is unlikely to be much variation in the hooks deployed per set but the variation in catch per set (for a given species) can be very high.

The preliminary analyses in this report and the development of operating models to scale observer data up to the level of logbooks laid the foundation for the study (though ultimately the operating models were not important). From preliminary analyses it was clear that for target species over a large area and an annual timeframe that sampling 10% of the sets on each trip was going to provide good precision for total catch. This is because of the large number of sets that occur each year in broadscale areas and because target species are caught on a good proportion of the sets. Indeed, for some target species, a sampling proportion of 5% was more than adequate.

For bycatch species it was noted in this study that using the magnitude of a CV, to measure performance of an estimator, was not always appropriate. For example, when there is a catch of 20 animals for some species it is not necessary to have a CV of 10% or less (e.g., a 95% CI of about 16-24). It is generally of little consequence whether there were 10, 20, or even 80 animals caught. Depending on a manager's tolerance for precision, for infrequently caught species (5-20% of sets), a sampling proportion of 10-20% will generally be adequate (see Figures 43 & 44). For the range of species looked at for longline over shorter timeframes and smaller areas, a sampling proportion of 20% was also generally adequate (e.g., to estimate catch at a regional level for bycatch species as well as target species). For purse seine catches a sampling proportion of 20% was also generally adequate (see Figure 42).

For rarely caught species (less than 5% of the sets) the main consideration may be whether the species was caught or not. In this case, a manager must consider the number of sets involved and the proportion of sets on which the species is generally caught. Then, given a threshold for the probability of incorrectly reporting a zero catch, they can calculate a required sampling proportion (or estimate it using one of Figures 45-48). For very rarely caught species it may be that to avoid falsely reporting a zero catch that most of the sets need to be sampled. Hopefully, in these cases the number of sets that need to be sampled can be reduced by restricting the area to the known habitat of the species.

It is recommended that:

- EM analysts randomly select at least 10% of sets from all trips for analysis
- That increased sampling proportions be used when good precision at a regional level is required:
    - By default, 20% of sets should be sampled at random from each trip
    - By design, a higher proportion of sets may need to be sampled to meet specific criteria (e.g., to determine if a species of special interest has been caught or not).

It is also recommended that the procedure and fields for recording the number of hooks between floats (hkbf), and the hook number on which an animal is caught be reviewed. It may be that more than one field is required per set for hkbf and/or that the hook number on which an animal is caught may need to be paired with the number of hooks in the given basket.

References

FFA, 2020. Regional longline fisheries electronic monitoring policy. Version 1 (adopted). 19 June 2020.

Lawson, T., 2003. Observer coverage rates and the accuracy and reliability of estimates of CPUE for offshore longline fleets targeting south Pacific albacore. SCTB16 Working Paper SWG-4.

Lawson, T., 2004. Observer coverage rates and reliability of CPUE estimates for offshore longliners in tropical waters of the Western and Central Pacific Ocean. SCTB17 Working Paper SWG-4.

Peatman T.; Nicol S., 2020. Updated longline bycatch estimates in the WCPO. WCPFC-SC16-2020/ST-IP-11

SPC, 2021. Longline Observer Guide, 2021: revised by Siosifa Fukofuka, Timothy Park and Caroline Sanchez.

SPC, FFA, PNAO, 2020. Report of the DCC meeting for the review of longline electronic monitoring (EM) data fields. WCPFC-SC16-2020/ST IP-07

Appendix 1: Equations for three simple sampling schemes

The statistical setting is sampling at random without replacement from a finite population. This means that the individual samples are not independent, and this must be accounted for in the variance equations using the finite population correction.

Assume we wish to estimate total catch for a given species over a specified time frame and area. Note, that the equations will hold for any field where we wish to estimate a total across all vessels, trips, and sets.

Assume that there are N vessels and vessel i has N_i trips and that on the jth trip by vessel i, there are N_{ij} sets. Denote the catch for the kth set on the jth trip by vessel i by c_{ijk}.

Sampling sets

Suppose that we sample sets at random for every trip by every vessel. We attempt to sample the same proportion p for each trip but we always sample at least one set. Denote the number of samples from the jth trip by vessel i as n_{ij}.

Let X_{ijs} = sample s of the set catches for the jth trip by vessel i.

For each random sample we have:

$E(X_{ijs}) = \mu_{ij}$ and $Var(X_{ijs}) = \sigma_{ij}^2$

where μ_{ij} and σ_{ij}^2 are respectively the population mean and variance of the set catches on the jth trip by vessel i.

To estimate the total catch across all trips and vessels we form the estimator:

$$Y = \sum_{i=1}^{N} \sum_{j=1}^{N_i} \frac{N_{ij}}{n_{ij}} \sum_{s=1}^{n_{ij}} X_{ijs} = \sum_{ij} N_{ij} \bar{X}_{ij}$$

That is, we use the samples from each trip to estimate the average catch per set for the trip. We scale up the average to estimate the total catch for the trip and then we sum over all trips. It is easy to show that Y is an unbiased estimator of the total catch:

$$E(Y) = \sum_{ij} N_{ij} E(\bar{X}_{ij}) = \sum_{ij} N_{ij} \mu_{ij} = \sum_{ijk} c_{ijk}$$

Also, using the finite population correction:

$$Var(Y) = \sum_{ij} N_{ij}^2 \frac{\sigma_{ij}^2}{n_{ij}} \frac{(N_{ij} - n_{ij})}{(N_{ij} - 1)}$$

where the sum is over the i, j: $N_{ij} > 1$.

When $N_{ij} = 1$ there is only one set and it is sampled and hence there is no variance contribution from that trip.

In this sampling scheme, the expected number of sets sampled is simply the total number of samples ($\sum_{ij} n_{ij}$).

Sampling trips within vessel

Suppose that we sample trips at random for each vessel and we record the total catch for each selected trip. We attempt to sample the same proportion *p* of trips for each vessel and we always sample at least one trip. Denote the number of trips sampled for the *i*th vessel as n_i.

Let X_{is} = sample *s* of the trip catches by vessel *i*.

For each random sample we have:

$E(X_{is}) = \mu_i$ and $Var(X_{is}) = \sigma_i^2$

where μ_i and σ_i^2 are respectively the population mean and variance of the trip catches for vessel *i*.

To estimate the total catch across all trips and vessels we form the estimator:

$$Y = \sum_{i=1}^{N} \frac{N_i}{n_i} \sum_{s=1}^{n_i} X_{is} = \sum_i N_i \bar{X}_i$$

That is, we use the samples for a vessel to estimate the average catch per trip. We scale up the average to estimate the total catch for the vessel and then we sum over all vessels. It is easy to show that *Y* is an unbiased estimator of the total catch:

$$E(Y) = \sum_i N_i E(\bar{X}_i) = \sum_i N_i \mu_i = \sum_{ijk} c_{ijk}$$

Also, using the finite population correction:

$$Var(Y) = \sum_i N_i^2 \frac{\sigma_i^2}{n_i} \frac{(N_i - n_i)}{(N_i - 1)}$$

where the sum is over the i: $N_i > 1$.

When $N_i = 1$ there is only one trip, and it is sampled and hence there is no variance contribution from that vessel.

To work out the expected number of sets sampled, the average number of sets per trip is calculated for each vessel and multiplied by the number of selected trips. These are the expected number of sets for each vessel which are then summed across vessels.

Sampling trips

In this scheme the trips are pooled across vessels and trip catches are sampled at random. Suppose that there are N_{trip} trips in total and that we sample n = round(pN_{trip}) trips at random.

Let X_i = the *i*th sample of the trip catches, then

$E(X_i) = \mu$ and $Var(X_i) = \sigma^2$

where μ and σ^2 are respectively the population mean and variance of the trip catches (pooled across vessels).

Let $Y = N_{trip}\bar{X}$ then *Y* is an unbiased estimator of the total catch and

$$Var(Y) = N_{trip}^2 \frac{\sigma^2}{n} \frac{\left(N_{trip} - n\right)}{\left(N_{trip} - 1\right)}$$

We can safely assume that we have more than one trip.

To work out the expected number of sampled sets the average number of sets per trip is calculated and multiplied by the number of samples (*n*).

Appendix 2: Details on generating an operating model

A summary of the approach has been given in the main text. This Appendix has some detail on the method used for tuning the variance of the trip catches while maintaining the consistency of the mean catch-per-trip with the mean catch-per-set of each trip.

In this context an operating model is created for each specified species and area. It contains everything that is needed to calculate the CV for each sampling scheme considered. It is built on the "skeleton" of the trips within the logbook data (everything except the catches). A predicted mean and variance for the set catches are filled in for each trip using models fitted to the observer data for the same species and area.

A binomial model is fitted to the observer data and then applied to the logbook skeleton to generate an appropriate number of positive and zero trip catches. For the positive trip catches, a log-linear model, that was fitted to the observer data to explain mean catch-per-set, is then used to generate the mean catch-per set for each trip in the logbook skeleton. This is done in two steps. First, the model predictions are generated. **Then they are modified by adding variance in such a way that the mean catch-per-trip is unaltered and the variance of the (positive) trip catches approximates the variance of the (positive) trip catches in the observer data.** Finally, the log-linear model that was used to explain the variance of catch-per-set for each trip (with positive catch) in the observer data is used to generate variances for the logbook skeleton.

The bold sentence in the paragraph above is explained further below.

The log-linear model (mean.mod) for mean catch-per-set is of the form:

log(mean) ~ year + other factors.

To generate the predictions in R a call to predict() was used:

pred <- predict(mean.mod, Logdata, type="response")

where Logdata is the logbook data which has positive trip catches (it contains the same type of factors/variables for each trip that were used when fitting the observer data). Note that the predictions are of log(mean) and have to be exponentiated to be on the linear scale.

The trip catches implied by the predictions are:

Trip.catch = Logdata$n × exp(pred + Y)

where Logdata$n are the number of sets for each trip and Y is a vector of random residuals from the model fit (to introduce stochasticity into the generation of the mean catch-per-set values). We take each $Y_i \sim N(0, s^2)$ where s is the standard deviation of the model residuals.

Denote the population variance of the elements of Trip.catch as Var(Trip.catch). Also, denote the population variance of the positive trip catches from the observer data as Var(Obs.trip.catch). We want to tune pred so that Var(Trip.catch) ≈ Var(Obs.trip.catch) without changing the expected value of exp(pred + Y).

This is done by adding a $N(0, \sigma^2)$ random variable to each element of pred and correcting for the bias induced by the exponential transformation.

That is,

Newpred = exp(pred + X) × correction

where X is a vector of the same length as pred and each $X_i \sim N(0, \sigma^2)$.

The correction factor is:

correction = exp(-0.5 × [$\sigma^2 - s^2$]).

Of course, we need a method to determine an appropriate value for σ^2. This was done by calculating a starting value and then implementing a simple search algorithm that updated σ^2 until Var(Trip.catch) ≈ Var(Obs.trip.catch) (within 10%).

For the k[th] trip with a positive catch, let

$$D_k = N_k e^{P_k + X_k} = N_k e^{P_k} e^{X_k} = T_k e^{X_k}$$

where

N_k = number of sets on the k[th] trip
P_k = predicted log(mean catch-per-set)
T_k = predicted catch per trip (no stochasticity)
$X_k \sim N(0, \sigma^2)$

The starting value of σ^2 was calculated by equating the expected value of the population variance of the D_k with the population variance of the observer trip catches (V_{obs}):

$$E\left[\frac{1}{N}\sum_{k=1}^{N}(D_k - \bar{D})^2\right] = V_{obs}$$

Noting that,

$$E(D_k) = T_k e^{\sigma^2/2}$$

$$Var(D_k) = T_k^2 e^{\sigma^2}\left(e^{\sigma^2} - 1\right)$$

it can be shown that the equation is a quadratic in $x = e^{\sigma^2}$.

It then follows that,

$$\sigma = \sqrt{\log x}$$

and

$$x = \frac{b + \sqrt{b^2 + 4aV_{obs}}}{2a}$$

where,

$$a = \left(\frac{N-1}{N}\right)\overline{T^2}$$

$$b = \bar{T}^2 - \frac{\overline{T^2}}{N}$$

$$\overline{T^2} = \frac{1}{N}\sum_{k=1}^{N} T_k^2$$

It can be that the variance of the predicted positive trip catches, without any stochasticity, is higher than the variance of the positive observer trip catches. In this case, the search algorithm stops because it reaches a maximum number of iterations without matching the two variances if they differ by more than 10%. The difference in variances might be valid or it could be an indication that the mean model is not appropriate (perhaps over-fitting the data which causes some predicted trip catches to be unrealistic).

Appendix 3: General equations

Below we derive the equations for two general cases. The first case covers random selection of sets for a randomly selected trip from a randomly selected vessel. The second case is where vessels are ignored and sets are randomly selected from a randomly selected trip. For both cases, the equations are derived using the general result given below.

Let,

$$T = \sum_{i=1}^{n} T_i$$

where each T_i is a random selection, **without** replacement, from a set of independent random variables $\{X_1, \ldots, X_N\}$ where $E(X_i) = \mu_i$, $Var(X_i) = \sigma_i^2$.

Then,

$$\mathrm{E}(T) = n\bar{\mu}$$

$$\mathrm{Var}(T) = n\left[\overline{\sigma^2} + \left(\frac{N-n}{N-1}\right)\mathrm{Var}(\mu)\right]$$

where, $\bar{\mu}$ and $\mathrm{Var}(\mu)$ are respectively the population mean and variance of $\{\mu_1, \ldots, \mu_N\}$ and $\overline{\sigma^2} = \frac{1}{N}\sum_{i=1}^{N}\sigma_i^2$.

Note, that this general result reduces to the finite population correction if the X_i are constants.

The proof of the result is straightforward once it is noted that because of symmetry the T_i are identically distributed so that $\forall i$ $\mathrm{E}(T_i) = \mathrm{E}(T_1)$, $\mathrm{E}(T_i^2) = \mathrm{E}(T_1^2)$ and $\forall\ i \neq j$ $\mathrm{E}(T_i T_j) = \mathrm{E}(T_1 T_2)$.

Also, from conditional probability theory,

$$\mathrm{E}(T_1) = \mathrm{E}_{\boldsymbol{X}}[\mathrm{E}(T_1|\boldsymbol{X})]$$

$$\mathrm{E}(T_1^2) = \mathrm{E}_{\boldsymbol{X}}[\mathrm{E}(T_1^2|\boldsymbol{X})]$$

$$\mathrm{E}(T_1T_2) = \mathrm{E}_{\boldsymbol{X}}[\mathrm{E}(T_1T_2|\boldsymbol{X})] \text{ where } \boldsymbol{X} = (X_1, \ldots, X_N).$$

Since the X_i are selected at random on the first draw, they each have probability $1/N$ of being selected, therefore,

$$\mathrm{E}(T_1) = \mathrm{E}\left(\sum_{i=1}^{N}\frac{1}{N}X_i\right) = \frac{1}{N}\sum_{i=1}^{N}\mathrm{E}(X_i) = \bar{\mu}$$

and $\mathrm{E}(T) = n\bar{\mu}$.

Also,

$$\mathrm{E}(T_1^2) = \mathrm{E}\left(\sum_{i=1}^{N}\frac{1}{N}X_i^2\right) = \frac{1}{N}\sum_{i=1}^{N}\mathrm{E}(X_i^2) = \frac{1}{N}\sum_{i=1}^{N}(\sigma_i^2 + \mu_i^2) = \overline{\sigma^2} + \overline{\mu^2}$$

Since, the remaining X_i are selected at random on the second draw, they each have probability $1/(N-1)$ of being selected, therefore,

$$\mathrm{E}(T_1T_2) = \mathrm{E}\left(\sum_{i\neq j}\frac{1}{N(N-1)}X_iX_j\right) = \frac{1}{N(N-1)}\sum_{i\neq j}\mu_i\mu_j$$

and, since,

$$\bar{\mu}^2 = \left[\frac{1}{N}\sum_{i=1}^{N}\mu_i\right]^2 = \frac{1}{N^2}\left[\sum_{i=1}^{N}\mu_i^2 + \sum_{i\neq j}\mu_i\mu_j\right]$$

it follows that,

$$\mathrm{E}(T_1T_2) = \frac{N^2\bar{\mu}^2 - N\overline{\mu^2}}{N(N-1)} = \frac{N\bar{\mu}^2 - \overline{\mu^2}}{N-1}.$$

Therefore,

$$\mathrm{E}(T^2) = \mathrm{E}\left[\sum_{i=1}^{n}T_i^2 + \sum_{i\neq j}T_iT_j\right] = n\mathrm{E}(T_1^2) + n(n-1)\mathrm{E}(T_1T_2) = n\left(\overline{\sigma^2} + \overline{\mu^2}\right) + \frac{n(n-1)\left(N\bar{\mu}^2 - \overline{\mu^2}\right)}{N-1}.$$

Noting that Var(T) = E(T^2) – E(T)^2, the result follows after some algebra.

Random sets from a random trip from a random vessel

From N vessels, n are selected at random without replacement. Given that the ith vessel is selected, it has N_i trips of which n_i are selected at random without replacement. Given that the jth trip of the ith vessel is selected, it has N_{ij} sets of which n_{ij} are selected at random without replacement. The estimator of the total catch for the whole fleet for a given species in a given area is calculated in a number of steps. First, the mean catch-per-set is obtained for each selected trip for each selected vessel. This is then scaled to give an estimate of the trip catch for each selected trip and vessel. The average trip catches for each selected vessel are then scaled to give an estimate of the total catch for each selected vessel. The average vessel catches are then scaled to give an estimate of the total catch for all vessels.

Denote the catch for the kth set of the jth trip by the ith vessel as c_{ijk} and denote the mean and variance of the set catches as $\bar{c}_{ij}$ and σ_{ij}^2.

Given the ith vessel and the jth trip, the average catch-per-set of the samples is

$$\bar{X}_{ij} = \frac{1}{n_{ij}}\sum_{k=1}^{n_{ij}}X_{ijk}$$

And it follows almost immediately from the finite population correction for sampling without replacement that,

$$\mathrm{E}(\bar{X}_{ij}) = \bar{c}_{ij}$$

$$\mathrm{Var}(\bar{X}_{ij}) = \frac{\sigma_{ij}^2}{n_{ij}}\frac{(N_{ij}-n_{ij})}{(N_{ij}-1)} \text{ for } N_{ij} > 1$$

with a zero variance if N_{ij} = 1.

The average catch-per-trip for the ith vessel is,

$$\overline{Y}_i = \frac{1}{n_i}\sum_{s=1}^{n_i} N_{is}\overline{X}_{is}$$

If we let $T_s = N_{is}\overline{X}_{is}$ then we can apply the general result to get the expectation and variance of each $\overline{Y}_i$. For given i and j, let:

$$c_{ij} = \mathrm{E}\left(N_{ij}\overline{X}_{ij}\right) = \sum_{k=1}^{N_{ij}} c_{ijk}$$

and let,

$$v_{ij} = \mathrm{Var}\left(N_{ij}\overline{X}_{ij}\right) = \frac{N_{ij}^2\sigma_{ij}^2}{n_{ij}}\frac{\left(N_{ij} - n_{ij}\right)}{\left(N_{ij} - 1\right)}$$

when $N_{ij} > 1$ and $v_{ij} = 0$ when $N_{ij} = 1$.

Let $\sigma_{i,trip}^2$ equal the population variance of the c_{ij} and $\overline{v}_i$ equal the mean of the v_{ij} then,

$$\mathrm{E}\left(\overline{Y}_i\right) = \overline{c}_i$$

$$\mathrm{Var}\left(\overline{Y}_i\right) = \frac{1}{n_i}\left[\overline{v}_i + \frac{(N_i - n_i)}{(N_i - 1)}\sigma_{i,trip}^2\right]$$

Finally, we have the estimator of the total catch over all vessels:

$$T = \frac{N}{n}\sum_{r=1}^{n} N_r\overline{Y}_r$$

Applying the general formula again, let:

$$c_i = \mathrm{E}\left(N_i\overline{Y}_i\right) = \sum_{jk} c_{ijk}$$

and let,

$$\sigma_i^2 = \mathrm{Var}\left(N_i\overline{Y}_i\right) = \frac{N_i^2}{n_i}\left[\overline{v}_i + \frac{(N_i - n_i)}{(N_i - 1)}\sigma_{i,trip}^2\right]$$

when $N_i > 1$ and $\sigma_i^2 = 0$ when $N_i = 1$.

Let σ_{vess}^2 equal the population variance of the c_i and $\overline{\sigma^2}$ equal the mean of the σ_i^2 then,

$$\mathrm{E}(T) = \sum_{ijk} c_{ijk}$$

$$\mathrm{Var}(T) = \frac{N^2}{n}\left[\overline{\sigma^2} + \frac{(N - n)}{(N - 1)}\sigma_{vess}^2\right]$$

We also need the expected number of sets that will be sampled in this case.

The number of sets sampled is a random variable:

$$S = \sum_{r=1}^{n} \sum_{s=1}^{n_r} n_{rs} = \sum_{r=1}^{n} S_r$$

The general formula can be used again and we get,

$$\mathrm{E}(S) = \frac{n}{N} \sum_{i=1}^{N} \frac{n_i}{N_i} \sum_{j=1}^{N_i} n_{ij}$$

Random sets from a random trip

From N_{trip} trips (irrespective of the vessel), n_{trip} are selected at random without replacement. Given that the selected trip is the *j*th trip from the *i*th vessel, n_{ij} sets are selected at random without replacement and the mean catch-per-set from the sampled sets is:

$$\overline{X}_{ij} = \frac{1}{n_{ij}} \sum_{k=1}^{n_{ij}} X_{ijk}$$

We form the estimator:

$$Y = \frac{N_{trip}}{n_{trip}} \sum_{s=1}^{n_{trip}} N_{s(ij)} \overline{X}_{s(ij)}$$

where $s(ij)$ denotes a random sample of one of the i, j pairs that index the trips across all vessels.

As before we can apply the general formula:

$$c_{ij} = \mathrm{E}(N_{ij} \overline{X}_{ij})$$

is the trip catch for the *j*th trip on the *i*th vessel.

Also,

$$v_{ij} = \mathrm{Var}(N_{ij} \overline{X}_{ij}) = \frac{N_{ij}^2 \sigma_{ij}^2}{n_{ij}} \frac{(N_{ij} - n_{ij})}{(N_{ij} - 1)}$$

when $N_{ij} > 1$ and $v_{ij} = 0$ when $N_{ij} = 1$.

Let σ_{trip}^2 equal the population variance of the c_{ij} and $\overline{v}_{trip}$ equal the mean of the v_{ij} then,

$$\mathrm{E}(Y) = \sum_{ijk} c_{ijk}$$

$$\mathrm{Var}(Y) = \frac{N_{trip}^2}{n_{trip}} \left[\overline{v}_{trip} + \frac{(N_{trip} - n_{trip})}{(N_{trip} - 1)} \sigma_{trip}^2 \right]$$

The expected number of sets sampled is the number of trips sampled multiplied by the average number of sets per trip (across all trips).