

4.2 Sampling catches, effort, CPUE and size

In principle, there are two data collection methods: complete enumeration, and sampling. A frame survey or fishery census is generally conducted by complete enumeration. For estimating the total annual catch, complete enumeration would be ideal. This is generally beyond the budget of most fisheries research centres, however. As a result, an appropriate sampling system is required to provide representative data that can be raised to the total fishery. A diagrammatical representation of the processes that can occur when assessing catch levels for Task 1 information is presented in **Figure 4.2.1**.

This section briefly considers statistical and related practical aspects of sampling tuna fisheries and fish in various ways in order to estimate summarising statistics for total landings, fishing effort, and size and other biological characteristics of the fish. Section 4.2.1 is a guide to basic sampling theory in the context of tuna fisheries. Section 4.2.2 describes and comments on common statistical sampling designs, and section 4.2.3 discusses sampling precision. Section 4.2.4 turns to more practical matters. It summarises the main different sources of information about tuna stocks and fisheries, drawing attention to their strengths and weaknesses, and suggests practical ways to sample and estimate in the light of the preceding statistical discussion. Section 4.2.5 presents potential problems in total annual catch statistics, and possible solutions. Finally, section 4.2.6 considers the raising of sample estimates to estimates for an entire fishing fleet or for a fish stock within a standard ICCAT time-area stratum, e.g. $5^{\circ}\times 5^{\circ}\times\text{month}$, $1^{\circ}\times 1^{\circ}\times\text{month}$.

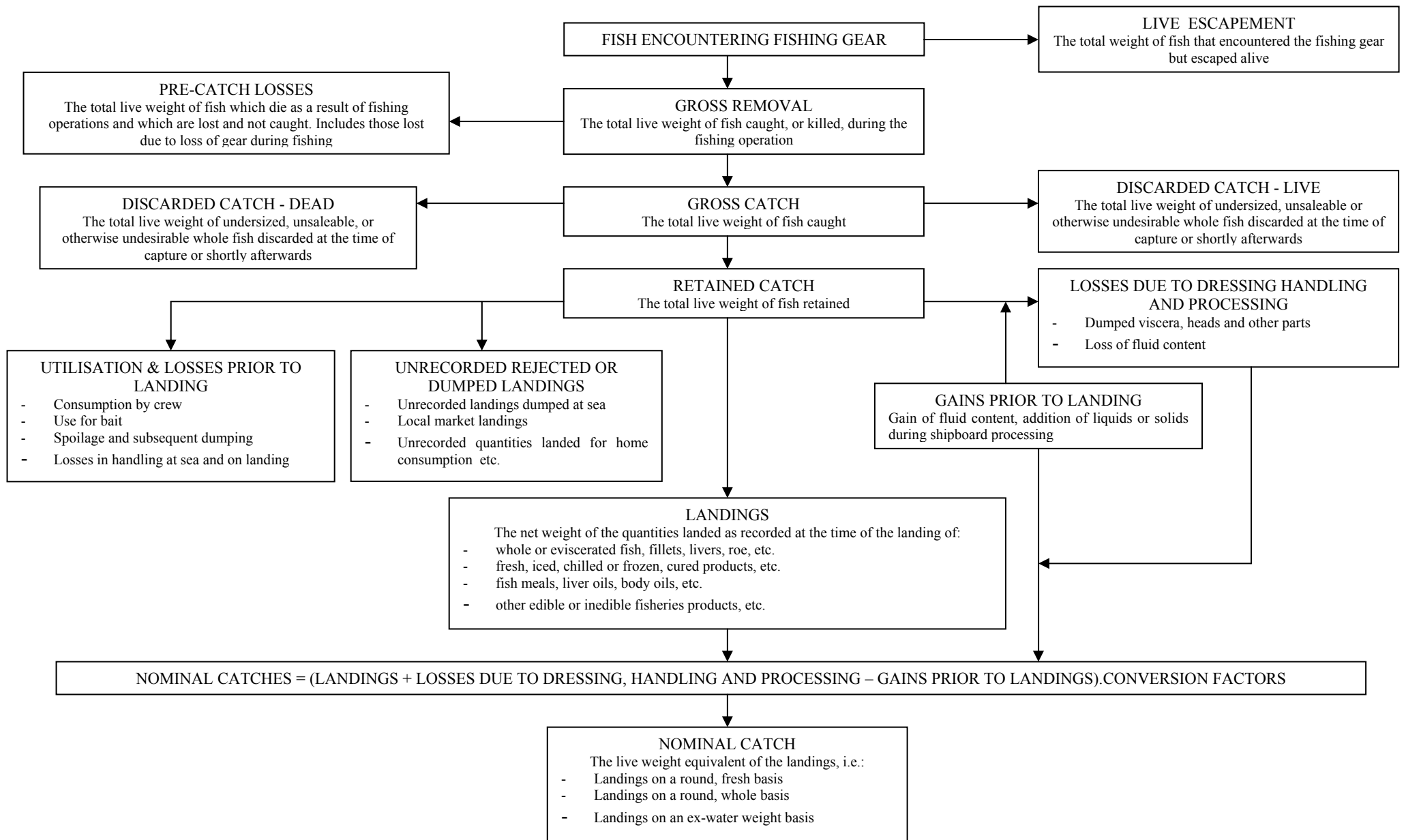


Figure 4.2.1. Diagrammatic representation of the concepts involved in the estimation of nominal catch

4.2.1 Sampling basics

Scientific sampling

Sampling is essential when we wish to describe or make inferences about a population that is too big to permit observation of every individual member. Scientific sampling requires that there be a relationship between the population and the sample. Two philosophies for this are (Thompson, 1992):

- Design-based sampling in which the relationship consists of a probabilistic rule for selecting individuals for observation; e.g. simple random sampling.
- Model-based sampling in which the population is a hypothetical construct based on a mathematical model whose parameters are estimated from the observed individuals; the model includes random errors, e ; e.g. $Age = f(Length) + e$.

Design-based sampling allows descriptive statistics such as the mean, variance, and frequency distributions to be estimated with no assumptions about the population. These statistics are *design-unbiased*, i.e. they are expected to fall around the true value with repeated sampling by virtue of the probabilistic sampling design. Model-based sampling, on the other hand, allows the model to be fitted with no assumptions about the sampling. The estimated parameters are *model-unbiased*, i.e. they are unbiased if the model is true and complete.

In practice, both philosophies require compromises because probability sampling can seldom be ideal, and no model is ever completely trusted (Burnham and Anderson, 2002). However, a random sample is usually suitable for fitting a model, whereas a sample taken to optimise the fit of a model is likely to be biased and inefficient for estimating descriptive statistics. Furthermore, sampling design is more or less within the control of the sampler depending on practical constraints, whereas incomplete knowledge of the important explanatory variables needed to model a fishery without bias probably cannot be rectified easily by a modeller. Modelling is an important research tool in fisheries science but a design-based approach based on the best practical approximation to a probability sampling scheme is advocated here for the collection of basic statistics about tuna fisheries.

The population

The population to be sampled and the nature of its individual members, called the *sampling units* (s.u.), should be carefully considered before designing a fishery sampling scheme because it is likely that the *observable population*¹ will be a subset of the *population of interest*. For studies of a fishing fleet, the population of interest may comprise all vessels in the fleet but the observable population may only consist of those vessels that are accessible in nearby ports. For biological studies, the population of interest may be all fish in the stock but the observable population may be restricted to the accessible part of the total landings from the stock. Since the population of interest cannot be sampled, a link between it and the observable population must be assumed. For fishing fleets, for example, assuming that the unobserved part of the fleet behaves in the same way as the observed part would provide the link and would require only a raising factor to convert survey estimates to estimates for the population of interest. For biological sampling, the observable landings, L , may be related to the stock, W , by a catchability function, q , of length, l , and fishing effort, E :

$$L_l = q(l).E.W_l$$

Such a model assumes that fishing is random with respect to the fish (Hilborn and Walters, 1992, p. 177). In both examples the assumptions are strong and may be controversial. The population of interest, the observable subset, and the assumed link between them should be stated explicitly in any sampling report. In the following text, the word ‘observable’ will be intended if not stated with ‘population’.

Randomisation

Statisticians call for random selection of sampling units so that estimation of parameters for the population can be justified by probability theory. In practice, this is often very difficult to achieve. The merits of striving for randomisation can be explained intuitively as follows. Consider the relatively simple task of estimating the average length of fish in the landings of a fishing vessel using simple random sampling. The constitution of the landings will have been affected by many influences such as the localities of the hauls, the season, the weather, the shift of crew that sorted the catches, etc. If the sample is collected from a restricted location in the bulk of the landings it may only reflect part of one or a few hauls taken in restricted circumstances, and the mean length of

¹ ‘Observable population’ is intended as a more self-explanatory term than ‘sampling frame’ of sampling theory.

that sample may, as a result, be very different from the (unknown) mean length for the bulk. In addition, the variability within a restricted sample is likely to be less than that in the bulk, meaning that the variance and confidence limits around the estimated mean are under-estimated, giving a spurious impression of good precision. By sampling fish as nearly as possible from randomly chosen locations in the bulk, the factors influencing the length of fish present can be expected to have the same proportionate effects on the sample. In statistical terms, randomisation provides (design-) unbiased estimates of the mean and variance. Achieving the best approximation to randomisation in different practical situations is discussed in section 4.2.4.

Information

Fishery sampling is usually costly so as much information should be taken and preserved from each sampling unit as possible. The information provided about the population by a single sampling unit depends on:

- The number of variables measured on the sampling unit, e.g. length, age, and maturity may all be measured on each fish drawn from a catch, although often only length is readily available.
- The relationships between those variables. e.g. maturity of fish is related to age. The observed values of different variables should be stored together and analysed as a vector for each sampling unit so that these relationships are preserved, e.g. for modelling. If not, information is lost.
- The precision of measurement. This is more of an issue for some variables than others. e.g. fishing effort is hard to measure on a certain type of vessel; maturity stage may be difficult to determine accurately for a fish. Ideally, the precision of measurements will be estimated for difficult cases by different people making independent measurements of the same variables on the same set of sampling units. A large measurement variance could negate the benefits of an elaborate sampling design carried out with much cost and effort.
- Whether the sampling unit was selected at random or by conscious choice. In the latter case, some of the information provided by the sampling unit relates to the method of choosing, i.e. to the bias, not to the population. An estimate is 'unbiased' if the average over many repetitions of sampling equals the true population value. Bias is not necessarily bad if it is constant and this is often a necessary assumption in fisheries work.
- The variance, σ^2 , of sampling units in the population; the information provided by a sampling unit is proportional to $1/\sigma^2$. σ^2 is estimated by the sample variance, s^2 .

The information provided about the population by a sample of $n > 1$ sampling units depends on:

- Whether the sampling units were drawn independently or in a pattern. The variance of the sample mean is estimated by s^2/n where n is the number of sampling units in the sample, but this is only true if the sampling units were drawn independently. Sampling units that are close together in space or time tend to be more similar than sampling units that are distant. Thus sampling units that are collected in a pattern will carry information that relates to the pattern and not to the population. For the same reason, sampling units drawn from compartments of a population are more likely to be similar within the compartments than among different compartments. Compartmentalised populations are common, e.g. fish arranged in a market by vessel, possibly also by size category; vessels using a particular port; fishing trips within a quarter. Statistical sampling schemes such as stratified, and multistage sampling are designed to isolate the variance among compartments; the sampling units are located independently and randomly within them. [Modellers would use mixed models to estimate the variances among compartments (Pinheiro and Bates, 2000).]

The number, n , of sampling units to include in a sample is usually a matter of taking as large an n as possible with the available sampling staff and resources. A more scientific approach requires decisions about the minimum acceptable precision, and the confidence level needed for a statement about whether the precision was achieved. These values may then be applied to sample size formulae, e.g. Thompson (1992, chapter 4; see box below) assuming that all sampling units are independently drawn.

Estimating sample size (based on Thompson (1992))

A population parameter θ (e.g. population mean) is to be estimated using an estimator $\bar{\theta}$. The aim is for the estimate to be close to the true value with high probability.

If the estimator $\bar{\theta}$ is an unbiased normally distributed estimator of θ , then $\frac{\bar{\theta} - \theta}{\sqrt{\text{var}(\bar{\theta})}}$ has a normal distribution. Using z to denote the upper $\alpha/2$ point of the standard normal distribution,

$$P\left(\frac{|\bar{\theta} - \theta|}{\sqrt{\text{var}(\bar{\theta})}} > z\right) = P(|\bar{\theta} - \theta| > z\sqrt{\text{var}(\bar{\theta})}) = \alpha$$

Variance of the estimator $\bar{\theta}$ decreases with increasing sample size n , so if the sample size is increased far enough, $z\sqrt{\text{var}(\bar{\theta})} \leq d$, where d is the maximum allowable difference between the estimate and the true value.

When estimating a population mean with simple random sampling, the sample mean \bar{y} is an unbiased estimator of the population mean μ with variance $\text{var}(\bar{y}) = \frac{(N-n)\sigma^2}{Nn}$, where N is the population size, n is the sample size and σ^2 the population variance. Setting

$$z\sqrt{\left(\frac{N-n}{N}\right)\frac{\sigma^2}{n}} = d$$

and solving for n gives the necessary sample size:

$$n = \frac{1}{\left(\frac{d^2}{z^2\sigma^2} + \frac{1}{N}\right)} = \frac{1}{\frac{1}{n_0} + \frac{1}{N}}$$

where

$$n_0 = \frac{z^2\sigma^2}{d^2}$$

If the population size N is large relative to the sample size n , so that the finite population correction factor can be ignored, the formula for sample size simplifies to n_0 .

For sampling designs more complex than simple random sampling, sample size can generally be selected in the same way, with sample size determined so that the half-width of the confidence interval equals the specified distance.

As an example, we try to identify the sample size needed to estimate the mean length of a large tuna population within 1cm of the true mean with 95% confidence ($\alpha=0.05$). The variance from previous samples thought to be representative of the population distribution is 52.2cm. Given the large population size, the formula for n_0 can be used. This gives:

$$n_0 = \frac{(1.960^2) * 52.2}{1^2} = 200.53 \sim 201$$

where the constant 1.960 is the upper $\alpha=0.025$ point of the standard normal distribution.

If relative error (r), the difference between the estimate and the true value, divided by the true value, is of interest, then the following criterion must be met:

$$P\left(\left|\frac{\bar{\theta} - \theta}{\theta}\right| > r\right) < \alpha$$

To estimate the population mean μ to within $r\mu$ of the true value with probability $1-\alpha$, the sample size formula is:

$$n = \frac{1}{\left(\frac{r^2 \mu^2}{z^2 \sigma^2} + \frac{1}{N}\right)}$$

If γ represents the coefficient of variation for the population, i.e. $\gamma = \frac{\sigma}{\mu}$, the sample size formula can be written:

$$n = \frac{1}{\left(\frac{r^2}{z^2 \gamma^2} + \frac{1}{N}\right)}$$

Therefore, the coefficient of variation is the population quantity on which sample size depends when relative precision is to be controlled.

Unfortunately, customers for fisheries data tend to be unwilling to specify precision and confidence less than “as good as possible”, so the pragmatic approach – which is incidentally much less work – is likely to be most popular, especially as few important fisheries are over-sampled. There is however a more important question which is how sampling resources should be directed among different sources of information, e.g. logbooks, landings, and catch sampling at sea. This depends on relative costs and available precision, the uses to which the combined data are put, and the independence maintained in the different data sets. No easy formulae are available to solve the problem generically and each case will probably need a research project. An example of such a project was by Beare *et al.* (2002) who used bootstrapping with different levels of random error to assess the influence of abundance indices on a stock assessment.

4.2.2 Sampling designs

Thompson (1992) describes a wide variety of statistical sampling designs and the estimation formulae to go with them. Other good sampling texts are by Cochran (1977), Raj (1968), and Sukhatme and Sukhatme (1970). The focus here is to introduce sampling designs likely to be most readily suitable for fisheries work. The final subsection comments on the optimisation of survey efficiency.

Simple random sampling without replacement

Simple random sampling (srs) means that each sample of n different sampling units (s.u.) drawn from a population of N has an equal probability of selection. A table of random numbers, or, in many circumstances, the best practicable simulation of one, is used to pick sampling units ‘without replacement’, i.e. so that no unit can occur more than once in a sample. Simple random sampling may be used on whole populations, on pre-defined subsets of them, e.g. within sampling strata, or within the hierarchical stages of a multi-stage sampling scheme (see below). Simple random sampling is a sensible sampling scheme when there is no prior information about the likely values of the variable of interest on different sampling units. If such information does exist and is known to be reliable, it can be used to design other sampling schemes that would deliver better precision per observation. Simple random sampling is satisfactory when you have no interest in variation of the variable over time or space. It has the valuable advantages of being simple to implement and to estimate from. Simple random samples are often adequate for fitting models, or for ‘post-stratification’ after sampling if there is a need to look at results by some variable such as e.g. age or sex.

Simple random sampling with replacement

Simple random sampling ‘with replacement’ means that an individual sampling unit is replaced in the population whenever it is drawn for the sample. Thus a sampling unit may occur more than once in any sample. Under simple random sampling with replacement, each possible sequence of n sampling units has equal probability. Ordinarily, this would be inefficient but it is useful when the possibility of repeated observations on one individual sampling unit is needed, e.g. to draw multiple trips on one vessel for observation.

Sampling with probability proportional to size

Sometimes prior information indicates that certain ‘large’ sampling units are likely to display a larger value of the variable of interest than ‘small’ sampling units. E.g. large fishing vessels may be expected to catch more fish than small. With sampling with probability proportional to size (pps) every sampling unit is assigned a probability of drawing that is proportional to the expected value of the variable of interest. A simple way to draw a probability proportional to size sample is to form a list of all sampling units in the population together with their assigned probabilities. Cumulative probabilities from zero to one are put in an additional column. Uniform random numbers from 0 to 1 are drawn and matched to the cumulative probabilities to find the next selected sampling unit. Probability proportional to size sampling is more efficient than simple random sampling, possibly much more, but only if the prior information used to assign selection probabilities is reliable. Otherwise it can be worse (Cotter *et al.*, 2002) and may not be worth risking. Special estimation formulae are needed to correct probability proportional to size for the sampling bias towards ‘large’ sampling units. Probability proportional to size is an alternative to stratification of sampling units by size and may be easier to implement.

Systematic sampling

In this scheme, there is a fixed distance or number of sampling units between each sampling unit selected for the sample, i.e. sampling locations are arranged in a one- or higher-dimensional grid. There have been many applications in marine science when a spread of observational effort over space and/or time is wanted, e.g. taking a measurement every x hours, surveying fish over a 2-dimensional grid of geographic points, etc. Estimation of means and variances from systematic samples often uses the formulae for simple random sampling without replacement but there are risks of bias in doing this due to the pattern of the sample:

- Trends or oscillations in the variable of interest may mean that the sampling grid finds more high values than low, or vice versa.
- Periodicity in the variable of a wavelength comparable to the grid interval may mean that most of the observations fall at the high or low part of the oscillations. This type of bias is called ‘aliasing’.

The starting point and orientation of the grid should be chosen randomly within the region of interest if that is feasible. Systematic samples are generally good for modelling over the time or spatial dimensions of a grid but another risk then arises. It is well known that arrangements of purely random numbers along a linear axis will often suggest the presence of trends (Kendall, 1976, para. 3.17) so any modelling of apparent trends over a grid should be supported by sound prior reasons for including the explanatory variables.

Stratified sampling

Samplers often compartmentalise a population into geographic, temporal, biological (e.g. length) or fishing mode (e.g. FADs versus free school) sampling ‘strata’. This can be for two reasons:

- To distribute observational effort evenly over space and time.
- To use prior information about variation of the variable of interest to improve the efficiency of the survey for estimating an overall mean and variance.

The first reason is common and valid for many practical purposes but is not necessarily statistically efficient for estimating a mean value. Samples are usually collected within each stratum by simple random sampling, or the best approximation to it. For best statistical efficiency, stratum boundaries should be located so that the within-stratum variances are as small as possible; i.e. locate the boundaries where discontinuities, or the steepest gradients occur in the variable of interest. Efficiency is also affected by the sample sizes allocated to each stratum. Proportional allocation assigns an equal number of observations per unit area or time. It is often a sensible choice when stratification is for the first reason. Optimal allocation assigns observations to each stratum in proportion to its size and its within-stratum standard deviation. This makes sense when stratification is for the second reason.

Stratified random sampling schemes require that at least two sampling units be located in every stratum in order that overall variance can be estimated. In practice though, 2 sampling units do not give a reliable estimate of a

within-stratum variance even if there are no missing values, and many more than 2 sampling units are preferable. There is a danger that a requirement to estimate a variance will lead to over-sampling of the least variable strata, causing expensive inefficiency. Strata can be thought of as a luxury: keep their number to a minimum when you do not have large resources for sampling, and settle for less geographic and/or temporal information. The estimates will then be more precise and dependable.

Multistage sampling

Fisheries scientists frequently encounter hierarchically structured populations; an example is fish within catches within fishing trips within vessels within a fleet. Such a population of fish could conceivably be sampled by simple random sampling ignoring the hierarchical structure but this is often impractical to arrange, e.g. if observers would have to transfer from one vessel to another at sea to sample different catches. A more convenient procedure is to draw a sample of 'primary' sampling unit (p.s.u.) at the highest level first, then a sample of 'secondary' sampling unit (s.s.u.) from each principle sampling unit drawn, then a sample of 'tertiary' sampling unit (t.s.u.) from each secondary sampling unit, and so on. For the example (and ignoring practical issues discussed below for observer surveys), one would first draw a sample of vessels (p.s.u.), then a sample of trips (s.s.u.) made by each drawn vessel, then a sample of catches (t.s.u.) from those fished on each drawn trip. Having sampled in this way, standard formulae are available for estimation of the mean and variance at each level. Theory also exists to examine the variance at each level and to adjust sample allocations among levels to improve efficiency. However, the improvements will not necessarily be practicable in a fisheries context.

Optimisation of sampling

Sampling is an expensive business, not just in fisheries, and much statistical attention has been given to maximising the quantity of information obtained per observation by variation of sampling designs, sample allocations, and formulae for estimation. Application of these ideas in a fisheries context can however be disappointing firstly because sample survey design is often tightly constrained by geographic and logistical factors, and secondly because fishery surveys tend to be concerned with more than one species. A survey that is optimal for one species could be hopelessly inefficient for another due to different geographic distributions etc. A multivariate approach may be possible using a principal component in place of the result for a single species to optimise the survey. However, when different species become relatively more important, there will be a high risk of finding a major lack of information. Ease and reliability of implementation tend to have more relevance to the design of fishery surveys than statistical efficiency.

Trials of intensive sample collections of either continuous large-size samples or repeated samples of smaller size are recommended to audit the quality of the sampling. Using this type of trial allows the assumption that the perception of the population is unaffected by the sampling regime.

4.2.3 Estimating precision

A precisely estimated statistic is here taken as one that falls closely around a fixed value on repeated sampling. An accurately estimated statistic is one that falls closely around the true value for the population on repeated sampling (see also section 4.1.1). Thus the sample variance of the statistic estimates precision; it estimates accuracy only if the sample and the estimation formula are unbiased. The mean square error is the sample variance plus the square of the bias. It estimates accuracy but since the bias is seldom estimable in fisheries work it is little used and the concept of precision measured as $1/(\text{sample variance})$ is of more use.

The sample variance is estimable in many cases from analytic formulae given in sampling texts. An important assumption is that every observation is made independently (c.f. section 4.2.1) which will usually be the case if sampling is randomised and in accordance with an established statistical sampling scheme. If not, the analytic formulae are likely to over-estimate sampling precision because dependence between the observations reduces the effective degrees of freedom used as the divisor in estimators of variance.

Most sampling formulae assume that the numbers of individuals in the sample and in the population are both known exactly. Consequently their variance (var) does not have to be taken into account when raising a mean for a sample to an estimate of the total for a population. If the raising factor is known exactly and x is a random variable, a basic result from mathematical statistics is that

$$\text{var}(kx) = k^2 \cdot \text{var}(x) .$$

In fisheries work, presumption of exact knowledge about the raising factor would sometimes be too optimistic. The corresponding estimation formula when k is also a random variable independent of x is (Goodman, 1960)

$$\text{var}(kx) = k^2 \cdot \text{var}(x) + x^2 \cdot \text{var}(k) - \text{var}(k) \cdot \text{var}(x).$$

An example of uncertain raising factors occurs in observer surveys when a single catch is sampled and the relative volumes of sample and catch must be estimated. The variance of the estimate for the whole catch would require the second estimator. More complicated formula must be used when k and x are not independent (Goodman, 1962; Bohrnstedt and Goldberger, 1969).

Analytic formulae may not be available, or may be difficult to use for statistics derived from a complicated estimation process; e.g. numbers-at-age based on an ALK which involves sampling for length, then sub-sampling for age (see section 4.3.6). A rough and ready approach to this problem is to compare the number and independence of observations contributing to each estimate. So, for example, the number of hard parts read for age might be used to gauge the relative precision of estimated numbers of fish at different ages particularly if the hard parts were collected on several fishing trips from different regions. However, for formal scientific work the bootstrap should be used.

The bootstrap is a computerised method that works by repeatedly sampling with replacement from an existing sample on the assumption that that sample is a good representation of the population. This is called re-sampling (Efron and Tibshirani, 1993; Davison and Hinkley, 1997). Bootstrapping an ALK created by double sampling is a two stage process: firstly the length sample is bootstrapped, then, for each re-sample created, an age sample is drawn from the available otoliths within each length class. The re-sampling should preferably follow the actual sampling process in all aspects. By repeating the re-sampling many times, a distribution is formed for the statistics of interest from which variances and confidence intervals can be estimated. Programming such a bootstrap can be complicated, and running it may take appreciable time, particularly for large samples. There is little point in undertaking the work if the original sample is thought to be seriously biased towards one locality, time period, or circumstance.

4.2.4 Information sources

Our ability to assess and forecast fish stocks and to understand migrations and processes regulating their success depends on our knowledge of the fisheries and the biology of the target species in different parts of the ocean. The total landings and total effort of a fishing fleet each year must often be estimated by a sampling procedure because collection of data from all fishing trips (a census) is too expensive or impractical. Biological information must also be obtained by sampling the fisheries because fishery-independent research vessel surveys at the scale of the Atlantic Ocean are likely to be uneconomic. This section describes three fishery-dependent methods for collecting information about the fisheries for, and biology of tuna-like species, namely logbooks, sampling of landings, and observer surveys. Other sources of information worth considering are canneries and sport fishing organisations.

The limitations of fishery-dependent information should be kept in mind (Paloheimo and Dickie, 1964; Hilborn and Walters, 1992; Swain and Sinclair, 1994; Rose and Kulka, 1999) because fishers tend to target known high concentrations of fish rather than acting as random samplers of the stock. They may therefore achieve high CPUE even though stocks are low. Furthermore, the catching power of individual vessels tends to increase over time as they are fitted with more powerful engines, and better fish finding equipment. The catching power of the fleet as a whole may also increase as a result of these changes, particularly if old vessels are replaced by new; alternatively, it may decrease due to a net loss of vessels for economic reasons.

Whether or not the following sampling procedures are adhered to, the sampling methods, formulae, and models actually used should be recorded in brief documents, here called 'Standard Operating Procedures' (SOPs). One individual can then take over a data gathering task from another without varying procedures, or, if procedures should be updated for some reason, a record can be kept of what changes are made and of when they occurred, as can be crucial for evaluating time-series. The SOPs should be available to ICCAT scientific groups so that the scientific value of the sampling results may be better assessed.

Logbooks:

Most fishing captains of purse seiners, longliners and baitboats will use logbooks to record the course of each voyage and the fish caught (see **Annex 1**). They may therefore be happy to use an adjusted design for the purpose of better monitoring the fishery, particularly if the design is drawn up in consultation with them and any others having an interest. A successful logbook scheme should provide low-cost information on quantities of fish retained for landing, fishing effort, landings per unit effort (LPUE), fishing strategies, and details of fishing

vessels. There may also be opportunities to collect additional useful data, e.g. on quantities of fish discarded and thus catch² per unit effort (CPUE). This is discussed further in section 4.4.

The success of a logbook scheme will partly depend on attitudes to it. The captain needs to understand fully the logbook and its purpose. A personal briefing and possibly some training, e.g. on species identification, may be needed. But it is also necessary for the fishery authority to take a constructive interest in all of the logbooks being completed since people soon lose interest in ignored communications. Regular examination and discussion with the captains should help prevent recurring errors or ambiguities, and regular feedback of summarised, useful information is likely to encourage continuing interest as well as serving as a check for errors. The feedback might be on a trip-by-trip basis or as a published annual report of aggregated data about the activities and landings of the whole fleet. Putting the logbook system on a laptop computer for use by the captain may be of mutual benefit if it reduces error-prone transcription work, and if the captain can then conveniently summarise or chart details of previous fishing activities.

A logbook system, however well designed, will be of low value if there are legal constraints on what the captain records, e.g. due to landings quota or restricted fishing areas. Promises that the data will be confidential may not be sustainable in a court of law and legal advice should be obtained before giving them. There may also be commercial restraints that inhibit a captain from declaring fishing positions and catches. Inadequate attention to accuracy and species identification are other potential problems for a logbook scheme, as is the omission of fish that are discarded or consumed at sea. In the absence of sampling data for unlanded fish, enquiries about policies on discarding, minimum acceptable sizes, and consumption of fish could greatly improve the scientific value of logbook data.

Installing a fishing logbook on every vessel in a fishery may not be practicable e.g. due to an unwillingness to co-operate, due to the geographic distance of a port, or due to the artisanal nature of the fishery. The possibility for bias then arises if, for some reason, vessels without logbooks fish differently from those that do. There may also be practical limits on the numbers of logbooks that can be collected and used. This, in contrast, is a sampling problem even if a large fraction of the fishing trips made are logged. Some randomisation of the vessels selected, with frequent changes, would help to diminish possibilities of biased inferences about the fleet as a whole that could result from repeatedly logging the same subset of vessels. Whatever the reason for incomplete logbook coverage of a fishery, raising the results from the observed vessels to the total fleet is needed to estimate total landings and effort (but not to estimate average LPUE). Estimation and raising are discussed in section 4.2.6.

Having decided that a logbook scheme is worthwhile, every logbook should record once **for each trip**:

- Identification of vessel, captain, owner.
- Details of vessel including type, flag nationality, gross registered tonnage, power of engines (preferably as transmitted to the propeller shaft, i.e. excluding power used for generators, refrigeration, winches, etc.); length (specifying whether length overall or registered length); capacity for fish; number of fishing crew and the times of any shifts worked (in case catch processing differs between them).
- Date, time, and port of departure and arrival including for stops during the trip.
- Time lost due to breakdowns, poor weather, or other interruptions.
- Details of any trans-shipments or landings of fish made during the trip.
- Specifications of fish finding equipment available on board;
- Generalised details of fishing gear, i.e. excluding modifications made from set to set. For nets, this would include mesh sizes (specifying whether knot to knot, or stretched mesh measurement), twine type and construction, and preferably a net plan. For long-lining the general details would include the total number of hooks, number of hooks between floats, hook type, and a general diagram of the dimensions of the longline.
- Generalised details of fishing techniques including shooting and hauling operations, typical fishing depths, immersion times, weather limitations on fishing.
- The target species for the trip, plus policies used by the crew to decide whether to discard or keep fish of different species, e.g. minimum landing sizes.
- The names of fish species that will be identified in the catch log if caught, and of those that are likely to be mixed up because they are difficult to separate or because the market does not require separation.

² 'Catch' is here understood to refer to retained + discarded fish. 'Landings' may be a subset of retained fish if some are consumed on board, trans-shipped, or subsequently discarded to make space for fish of higher value.

[Vernacular names are sometimes confused (e.g. bonito for skipjack) so species' identities should be checked and translated to Latin names for the logbook records archived by the fishery authority. See section 4.2.5.]

- Methods used to estimate quantities of fish retained (and discarded if possible).

Many of these details need only be copied from trip to trip.

Logbooks should record **for each day** during the trip (regardless of whether or not a catch was made) the date, noon position, position of fishing, activities, the times spent steaming, scouting for fish, and fishing, amount of fishing effort employed, and catch by species.

Logbooks should record **for each set**, whether or not it produced fish:

- Gear deployed (if it changes from set to set). Details given should be sufficient to calculate a useful measure of effective fishing effort for each set.
- Positions and times of shooting and hauling, plus way-points if the vessel did not travel directly between the two.
- Damage sustained by gear during fishing.
- Weather and sea state. Oceanographic variables by arrangement with the fishery authority, depending on the sensors available for use.
- The retained quantities of each species, and mix of species, as numbers or weights.
- And, if available, estimates of the discarded quantities of each species and mix.

Logbooks should record **at the end of each trip**:

- Total time in stated units (e.g. hours, working days, 24-hr periods, etc) spent looking for and catching fish.
- The total quantity landed as registered by a commercial scale, preferably separately for each species and mix of species.

At the end of a trip, a check should be made to see whether the total landed quantity recorded for commercial sale matches the total quantities retained from each set. A systematic estimation error could easily arise, particularly if weighing equipment is not used on board. If so, the logbook records of daily quantities retained should be adjusted proportionally so that their sum matches the total landed weight (less any quantities known to have been lost during the trip). The adjustment factor should be recorded with the data. Another way of improving the scientific value of logbook data after landing is to add more detail about the species compositions of each catch. This is possible if mixtures of species are separated and their weights recorded separately as part of the marketing process.

The days in which a vessel was searching for fish but unable to catch any must be considered as fishing days. It is often misunderstood that fishing days mean only the days that yield catches. Scouting for fish is one type of fishing activity. Therefore, the logbook must be designed and instructions added to report activities, i.e. what the fishing vessels were doing on days when no catches were made. Whether the vessel was floating due to bad weather or gear breakdown, was moving from one fishing ground to another, or was looking for a school makes a difference in counting fishing days. Another measurement of effort, widely used for small fishing vessels, is the "search time" - the amount of time per day that the boat is actively looking for fish. The "search time" is calculated by subtracting the cruising time from the "fishing day". These data can be collected from the logbook, directly by observers, or can be estimated from observer's data. Where FADs are used, new measurements of fishing effort have been introduced. These include the number of sets, number of sets obtaining a catch, and average catch size of the set. All of these should be noted by the type of fishing gear.

Effort data must be reported in number of hooks for longline and in fishing days for surface fisheries. The number of hooks between floats is also used as a unit of effort in the case of multi-species fisheries. If this is not practical, the unit of effort should be chosen to reflect the effort directly made to harvest the corresponding catch. Recommended units of effort are listed below in descending order of preference for each type of fishing gear.

Longline:

1. Total number of effective hooks used (excluding hooks which were not effective in fishing)
2. Total number of hooks used
3. Total number of sets of longlines
4. Total number of boat-days fishing
5. Total number of boat-days at sea (out of port)
6. Number of hooks between floats
7. Total number of trips (cruises) made
8. Total number of boats actually engaged in fishing

Pole and line (baitboat)

1. Total number of boat-days fishing (including searching days, whether or not fish were actually caught). The number of boat-days spent baiting should be excluded, but can be recorded separately for use in assessing bait stocks.
2. Total number of boat-days at sea
3. Total number of fishing poles used; that is, the number of crew engaged in fishing with a pole
4. Total number of trips (cruises) made
5. Total number of boats actually engaged in fishing

Purse seine, ring net, lift net (bag net), seine net, gill net, trawl

1. Total number of boat-days fishing (including all days, whether or not fish were actually caught)
2. Total number of boat-days at sea
3. Total number of searching days (excluding the time spent in setting and hauling the net)
4. Total number of trips (cruises) made
5. Total number of boats actually engaged in fishing

Trolling, handline

1. Total number of hook- (or line-) days fishing
2. Total number of boat-days fishing
3. Total number of boat-days at sea
4. Total number of trips (cruises) made
5. Total number of boats actually engaged in fishing

Traps

1. Total number of trap-days (trap units multiplied by days at sea)
2. Number of trap units in operation

FADs fishery

1. Total number of sets
2. Total number of positive sets

Landings

Sampling of landed fish may be necessary to estimate total landings by a fishing fleet. It can also provide useful information about length and age compositions, weight-at-length, maturity-at-length, and other biological characteristics of a stock although they may be restricted by a requirement to purchase fish that are dissected or otherwise damaged during observation. The present section discusses sampling of a population of landing events. Estimation and/or raising to the level of total fleet landings, or landings for a time-area stratum, is discussed in section 4.2.6.

There are a number of potential biases in examining landings (as opposed to catch), which need to be remembered. These include:

- Fish caught are generally kept until the end of the trip, and may be unloaded at a port far from where the catch was made (particularly in the case of industrialised fisheries). In this case, the area and time of landing may be quite different from those of the catch. For example, some catches in the Atlantic Ocean can be unloaded in Pacific or Indian Ocean ports the year following the year of the catch was taken.
- Fish may be processed to some extent on board the vessels (e.g. dressed, filleted, gilled-and-gutted, frozen, or even canned).
- Fish may be eaten at sea by the crew, or discarded.

The first consideration is whether sampling is directed at a fleet, a stock, or a time-area stratum. This defines the ‘population of interest’ (cf. section 4.2.1). In practice, only landed fish, or a subset of them, are accessible for sampling so they form the ‘observable population’. The models to be assumed to link the observable fish to the total landings and thence to the fish in the population of interest should be discussed and documented in a Standard Operating Procedure (SOP) before undertaking any expensive sampling. Possible linking models might be:

- ‘observable landings per unit effort are the same as the total landings per unit effort’. Effort can then be used as a raising factor.
- ‘All fish > X cm were retained’ implying that landings = catch above X cm.
- An elaborate model used to assess the stock from landings data.

Any biases in these assumed models will be added to biases in the sampling of the landings so it will be good to review them regularly and, if possible, to strive to minimise their importance, e.g. by seeking access to previously inaccessible landings, or by implementing an observer programme to estimate discarding (see also **Table 4.2.1**).

Landed fish are observable either in a fishing vessel before it has unloaded, on the quayside before transportation to buyers, or in a port market prior to sale. Given a choice, the best site to sample will be the one offering the sampler best access to the fish and most time to work before they are moved on. Observation at other sites, e.g. on a freighter used to trans-ship fish to port, possibly from more than one fishing vessel, may not be worthwhile if the origins of the fish are uncertain. Landings are a subset of total catch if any discarding takes place at sea, or if some fish were transhipped or consumed at sea. In the absence of sampling data, information on what happened to unlanded fish during a fishing trip should be sought if available.

The method of choosing which landing events to include in a sample should be decided in advance of a sampling programme so that it may be used consistently. A complication is that we do not know in advance when, or how many landings will occur. The timing of landings might be influenced by season, weather, day of the week, fishing location, identity of the captain, and a host of other potential factors. Systematically arriving at a port to sample on alternate Wednesdays, say, could result in bias from one or more of these factors. Randomising the sampling days within a sampling period, e.g. a quarter or year, offers less scope for bias except that vessels making short trips (and landing frequently) will be encountered more often than of those making long trips. As a result, inshore stocks would probably be better represented than distant water stocks in the samples. The identity of the population of interest (cf. 4.2.1) is important here. If it is ‘the total landings of the fleet’, then randomly choosing sampling days is a reasonable policy because short and long distance fishing trips are expected to be represented in the sample in the same proportions as they occur in the population. If, on the other hand, it is ‘the total stock of fish’, and much of the stock is thought to be present in distant waters, frequent occurrence of short trips in a sample would represent a bias. Stock-orientated sampling also suffers from the additional complications of confounding of gear and vessel effects with geographic location. For these reasons, it is probably an approach to avoid if possible.

Another option for randomly sampling landings by a fleet would be to pick vessels randomly with replacement from a list of the whole fleet, as suggested for observer surveys below, provided that practical arrangements can be made to meet specific vessels when they land. This option is more trouble to implement but would be better if landings of some vessels reside in port for appreciably longer than others, a factor that could bias sampling based on randomly timed visits. Stratification and probability proportional to size (cf. 4.1.2) based on vessel sizes or activities are further options for sampling landings. The population of interest and the sampling scheme should both be documented in a SOP.

Landings can be available for sampling in several different ways. The fish may be in bulk, e.g. as a pile, or as a mass of fish in the well of a vessel, or they may be arranged in boxes or other containers. There may be one species present or several mixed together. There also may be different size or freshness categories. The following notes are an attempt to cover most circumstances:

- Sampling bulk fish:

There is no guarantee that a large load or sampled well of fish will be homogeneously mixed; rather, the opposite is more likely. Ideally, sampling would give each fish in the load an equal chance of inclusion in the sample but this is seldom achievable in practice given restricted physical access, time, and other typical difficulties. So, to maximise the accuracy of the mean of the sample, the sample should be a composite from sub-samples of fish taken from several locations in the load, e.g. from the centre and all extremities, surface

and bottom. If this is impractical, the sampler should try to devise other practical ways to minimise his/her influence over the choice of individual fish, so that all possible classes are present in the sample in roughly the same proportions as they are in the bulk. Classes of fish that might easily be over-represented for human reasons are big, small, “representative”, eye-catching, and obscured-but-searched-for-anyway. To improve accuracy further, the number of fish collected in a sample should be large but only if the fish are selected independently. Time spent making many measurements on a large sample collected from just one location in the bulk would probably be better spent in trying to get fish from other locations and then accepting a smaller sample. Further comments on the numbers of fish to take in a sample are given below.

- Mixed species:

Fisheries data are seldom of much use unless they are linked to known species. When faced with a mix of species for sampling, the first basic task is to estimate the proportions of each. Often the mix will be present as a bulk load. In this case, the comments above on accurate sampling apply except that it will probably be necessary to take larger sub-samples of fish from each location in the bulk as a practical step to minimise the sub-conscious influences of the sampler on which species of fish are included. Each sub-sample should be of approximately similar size. The estimate of the proportion of species s in the load will then be the numbers, n_s , of individuals of that species in the composite sample divided by the number, N , of individuals of all species. The sample may be adequate for subsequently estimating biological characteristics of the common species but not for the rare species. Additional, monospecific samples are to be collected if required, again by looking for each species at different locations in the load. The issues of sampling multi-species catches are now discussed in more detail.

In some cases, the complexity of the catch requires a complicated sampling scheme. The purse seine fishery for tropical tuna falls into this category. These fisheries often represent a combination of species and fishing types, where the declaration of catch per species may also be strongly dependent on fish size. A description of the simultaneous sampling procedure for this particular fishery is given below as an example of combined sampling. This procedure allows the specific composition and the size distribution of the catch to be obtained.

In the Atlantic and Indian Oceans, sampling strategies for tropical tuna from purse seiners in port have been derived (Sarralde *et al.*, 2005). Multispecies catches are particularly common when fishing on FADs. Strata are based upon the geographic location of catches, time and association (e.g. FADs, free schools), all identified from the vessel logbook and well plan. Preferably, a well that contains fish from sets belonging to a single stratum (location, time and type of school) should be sampled. In exceptional cases, depending on the amount of sampling and forecasts made, wells in which sets not belonging to the same geographic zone or time stratum but close in position (less than 5° difference) or in time (fewer than 15 days difference) could be considered valid. However, sampling should **never** be undertaken on wells containing fish from different associations. Sampling should not be concentrated either in time (all months in a quarter should be sampled) or in space (all zones should be sampled). 15-25 samplings are recommended for each stratum.

Once priority wells for sampling have been selected, sampling can begin. It is recommended that at least two persons are present for sampling, one for selecting and measuring, and the other for noting the data on appropriate forms. A safe sampling site should be selected, ensuring that unloading is not hindered and that access to the fish is easy and safe. Sampling can occur where the fish are unloaded via the deck (with the agreement of the Captain), on the conveyor belt, or where larger fish and smaller fish are separated they may be measured separately, ensuring no pre-selection occurs.

Sampling different areas of the well (e.g. top versus bottom) may result in a different species composition. To avoid this, sampling should always be carried out on each well in two stages, or through sub-sampling. The first stage should be done shortly after the well is opened, and the second several hours later, but before well unloading is finished.

If unloading involves any selection of species (by species or weight category), the sampler must take the sample directly from the well. If there has been no selection, sampling may be performed during unloading, but always at random.

If the well only contains large fish (>70cm), 100 specimens should be measured at each stage (200 in total). All specimens (mixed species) should be taken at random until the optimum number is reached.

If the well only contains small fish (<70cm), during the first sampling stage 300 specimens should be taken (all species included). If this involves skipjack, frigate tuna or little tunny, the first 25 species per species can be measured, and the rest counted. If it involves bigeye, yellowfin or albacore, all specimens should be measured. During the second stage, 200 specimens should be measured and/or counted in a similar fashion.

If wells contain a mixture of large and small specimens, a total of 300 specimens should be measured and/or counted (all species included) at the first stage. If this involves skipjack, frigate tuna or little tunny, the first 25 specimens per species should be measured, and the rest counted. If this involves bigeye, yellowfin or albacore, all species should be measured until the recommended number is reached. At the second stage, 200 specimens should be measured and/or counted in a similar fashion. The weight of both weight categories in the well (fish above and below 10 kg) must be known. **Annex 2** includes the sampling manual of the tropical tunas for the Atlantic and Indian oceans.

Sampling of baitboats can be carried out following the same approach as listed for purse seiners above. The sampling unit in this case is the entire boat, rather than a single well. Generally, only one sampling will be performed, unless the vessel is large in which case two samples can be taken. When the catch is selected by size, species or commercial category before unloading, or if it is accessible to the samplers, a fraction of the fish will be sampled. As a result, all categories present will be randomly sampled. The number of categories into which the catch is divided defines the number of samples. The weight for each category should be noted.

As an example, fresh fish (from the most recent sets) may be unloaded from one side of the vessel, and frozen from another. Two samplings should therefore be made, one of the fresh fish, and one of the frozen. Similarly, weight should be known for both landings.

Catches from several vessels may be mixed together, for example where bait boats unload onto purse seiners or merchant ships. In these cases, no information is available on the fishing area or mode of fishing. As a result, a single sampling of the entire vessel must be performed.

- Containerised fish:

When landings have been containerised, e.g. in boxes, the containers themselves must be sampled. A random sampling scheme based on rows, or storage areas, is relatively easy to arrange using random numbers to pick the next container for measurements. Very large containers may themselves have to be sampled as a bulk load (see above), thereby creating an extra stage in a multi-stage sampling scheme (see 4.2.2). Containerised fish may be frozen, gutted, de-headed, etc. Their state should be reported and appropriate conversion factors applied to estimate the condition of the fish in life. Conversion factors that may be useful in tuna statistics are given in **Appendix 4**.

- Categorised fish:

Fish categorised into size or freshness grades will probably be containerised in some way. The containers should be sampled as described above. A sample that omits one or more categories evidently has a major potential for bias so the availability of all categories should be established before beginning observations. In addition, the total weights or volumes of each category in the landings must be known so that the results for each can be weighted appropriately in an estimate for the whole landing.

- Artisanal fisheries:

Sampling of catches from artisanal fisheries will largely take place at the landing site or market. Sampling at market will generally limit the accuracy of information on the fishing location, but artisanal fisheries are unlikely to operate too far from shore. Fishers can be interviewed to obtain desired information on fishing techniques and locations. Physical sampling of artisanal catches will merely require modification of the forms of sampling already described above. The unit of sampling will be at the scale of the vessel.

- Catch and release (sports) fisheries:

Catch and release fisheries provide an additional source of catch rate information that can prove important in particular countries. Methods to collect this information are described in Guthrie *et al.* (1991). Biological sampling of catch and release (sports) fisheries by definition can only occur while on board the vessel. Therefore, sampling will generally conform to that described for observers (see below). However, note that since samples are returned to the sea and mortality is negligible, there is a small but present danger of sampling individuals twice if an individual is subsequently caught in a commercial fishery. It is

recommended that sports fisheries be utilised for tagging and other biological studies (see sections 4.6 and 4.7).

Having obtained a sample as randomly as possible, the next task is often to estimate biological characteristics of interest. A length frequency distribution (LFD) is usually the first priority for each species (see also section 4.3.1). The number of fish that should be measured depends on how many modes (peaks) are present in the LFD. At small sizes, these will probably represent successive year-classes, although the size selectivity of the fishing method will affect their relative frequencies. Sufficient fish should be measured to define all of the modes present. In practice, this means measuring until modes are identified, then measuring, say, an extra one third of the existing sample to see whether any additional modes appear. Clear definition of all the modes is extremely helpful for distinguishing different age-classes. When there are clearly only one or two modes, a small sample may be sufficient, say 50 fish. When there are many modes, measurement of 300 or more fish is likely to be necessary. Adjusting the sample size in this way according to the results found requires that the initial random sample must have an excess of fish and be homogeneously mixed. Alternatively, additional samples of fish could be collected from the landings in exactly the same way as the first sample.

Some biological characteristics vary according to the size of fish. There are three options for estimating the relationship:

1. Estimate the characteristic on all fish in the original sample.
2. Remove a sub-sample of fish from each of a set of length classes and estimate the characteristic for these fish only³.
3. Remove a sub-sample without regard to length and fit a model.

Option (1) provides the biggest sample but there must be time and facilities to process every fish. Also, the most frequent sizes might be relatively over-sampled. For a statistical analysis of options (1) and (2) in connection with estimation of age composition, see Smith (1989).

Options (2) and (3) are suitable when sub-sampling is restricted. Option (2) is often used to develop age-length keys (see section 4.3.6, and Westheim and Ricker, 1978; Lai, 1993) using annually marked hard parts (see section 4.9). It could also be used to estimate maturity- or weight-at-length (see section 4.8). The sub-samples should be randomly selected from each length class but, in practice, provided that the characteristic being estimated is not visible to, and does not influence the sampler (as e.g. age and maturity would not), the first fishes that come to hand from each length class are suitable for the sample. Taking a fixed number of fish in the sub-sample from each length class is relatively easy to implement but may not be the most efficient practice (Kimura, 1977; Lai, 1993). A separate proportion-at-length or mean-at-length is estimated for each length class with option (2). There is no model or major assumptions but the large number of values to be estimated can result in low sampling precision, particularly if sub-sample sizes are small. The double sampling involved in this procedure complicates a statistical analysis to estimate standard errors.

Option (3) requires that care must be taken to ensure that the choice of fish for the sub-sample is not influenced by their size. Fitting a model is likely to require estimation of fewer parameters than estimation of mean values for many size classes (option 2) and the estimation can therefore be done with better precision. Disadvantages are that a model has to be assumed, and rare size groups are likely to be poorly represented in the sub-sample.

Observers

Observers travelling on fishing vessels can provide high quality data on quantities of most or all species retained and discarded, fishing effort, fishing methods, strategies and in some cases biological characteristics such as length frequency distributions (see section 4.10). As a spin-off, they can promote good communications between scientists and the industry. Observers may or may not have a role in enforcing fishery regulations. For purely scientific purposes it is better if they do not so that the captain fishes normally, without fear of prosecution. Observers may or may not have a right to travel on a fishing vessel. If not, the observable population of vessels may be restricted to those with co-operative captains and owners, leading to a possibility for bias. Observer programmes tend to be expensive because observers must be scientifically trained and must spend long periods at sea. Observers should be supervised at least on their first trip to make sure that they identify species and carry out all biological sampling competently. They should also be trained in marine, and specifically fishing safety (Luo *et al.*, 1999).

³ Thompson (1992, p143) refers to this as 'double sampling for stratification'.

An observer is mandatory on all vessels in a few fisheries. More commonly, observers must pick which vessels and trips they sail on to get most information about fishing and fish stocks. The 'population of interest' (cf. 4.2.1) is best defined in terms of the fleet of fishing vessels. Defining it as the stock of fish would require that trips be selected to evenly sample the different geographic areas occupied by the stock, a difficult task if the fleet focuses on certain parts of the stock, if vessels change fishing destinations during a trip, or if different gears are used in different areas causing a confounding of gear with geographic influences. Another complication is that the observable trips to be made by a vessel are often not known in advance due to interferences from poor weather, mechanical problems, poor fishing prospects, etc. Observer surveys are inherently hierarchical because sets (catches) are nested within trips within vessels but logistical constraints may prevent implementation of a multistage sampling scheme.

A first necessity for an observer sampling programme is a list of all vessels in the fishing fleet of interest. If necessary, obtaining a complete list would probably justify special research. The list should be updated before each sampling period to allow for changes in the fleet.

Secondly, information should be sought about the type, power and size of the different vessels, and about their history of activities, and landings or catches. This should be examined to decide whether there is sufficient reliable information for using a stratified or probability proportional to size sampling scheme based on an indicator of the fishing power of each vessel. Such a survey would be more efficient than simple random sampling but only if the information proves to be a good predictor of performance specifically for the forthcoming sampling period.

Thirdly, a decision is needed about whether to observe a sample of vessels (principle sampling unit) with several trips (secondary sampling unit) on each during the sampling period, or whether to try to observe fishing trips randomly from all those made by the fleet. The former would be a two-stage scheme that allows estimation of a between-trip-within-vessel variance; the latter is simple random sampling for trips. For large, diverse fleets and a relatively small number of observers, simple random sampling is probably the better choice because variation between trips on different vessels is likely to be larger than between trips on the same vessel; hence observation of as many vessels as possible is desirable. If there are sufficient observers to observe vessels repeatedly during a sampling period, or if only a few vessels within the fleet are observable, sampling of vessels and trips may be more informative if carried out as a two-stage scheme with all principle sampling units observed on more than one trip.

Simple random sampling for trips may be arranged by numbering all vessels in the fleet from 1 to V . Using random numbers from 1 to V , draw vessels from the list 'with replacement', i.e. so that the same vessel may occur more than once in the draw. Observers attempt to arrange trips on vessels in the order of drawing. Trip-level estimates can then be treated as independent observations of fishing activities by the fleet since each trip was independently selected and observed. Means and variances are computed from the trip-level estimates using the formulae for simple random sampling without replacement because, although vessels are selected with replacement, no trip is sampled more than once. This sampling scheme may be criticised if some vessels are known to spend many more days at sea than others per sampling period. In that case a more elaborate scheme could be used, e.g. probability proportional to size or stratification based on activity. Two-stage sampling could be arranged as for simple random sampling but fewer vessels would be drawn and more trips would be observed on each. Randomising the dates of trips on each vessel within the practical constraints is desirable.

An intuitive way to sample trips is to arrive at fishing ports on randomly chosen days during the sampling period and to choose the next vessel to go to sea. This scheme could be biased towards vessels spending most time in port and, possibly, to those offering the best conditions at sea for the observer. It is not recommended.

Once on board, the observer will collect information on the boat and fishing activity, and will pay special attention to data that cannot be estimated in any other way, such as discard estimation and identification and size measurement of accessory species. Regarding the target species, the observer should sample as many catches as practicable (without jeopardising safety through personal tiredness). Discarded and retained fish should be recorded separately. If only subset of catches can be sampled, the times of day should be varied and sampling should try to include catches from each different fishing locality visited. Each catch may itself be sampled in which case a raising factor to raise the sample to an estimate of the total catch must also be estimated. This might be achieved from relative volumes, or by relative times on a conveyor belt, etc. Sampling the catch should be guided by the procedures described above for sampling landings. The observer should record fishing effort for each set including those that are not sampled; less satisfactorily, at least a count of the sets not sampled should be

taken from the vessel log. Results for catches sampled may then be raised to estimate results for all catches taken during the trip.

4.2.5 Potential problems in total annual catch statistics

Table 4.2.1 presents typical problems relating to the collection of Task I statistics within ICCAT, and potential solutions.

Table 4.2.1 Problems relating to collection of Task I statistics.

| Problems | Examples | Solutions |
|--|---|---|
| <i>Species breakdown</i> | | |
| 1. Species reported together. The same prices per unit of weight paid. | <ul style="list-style-type: none"> - Portuguese Islands fisheries - Eastern Atlantic tropical surface fisheries (French/Ivorian/Senegalese fleet, Japan, Korea, E.S., etc. - Ex-Soviet purse seine fishery (all tunas combined) | <ul style="list-style-type: none"> - Encourage, educate, instruct and/or oblige fishermen to report catch by species - Carry out sampling and examine sample species composition in order to estimate catch species composition |
| 2. Misidentification or un-identifiable species. Lack of easy, clear identification key. | <ul style="list-style-type: none"> - Young yellowfin vs. bigeye tuna | <ul style="list-style-type: none"> - Find easy key and inform fishermen - Carry out sampling to estimate total species composition |
| 3. Confusion in local vernacular names. | <ul style="list-style-type: none"> - Spanish names for albacore (bonito), yellowfin (atún) - Japanese, Korean and Chinese names for marlins (black marlin locally called white marlin, blue marlin called black marlin) - Portuguese names for tunas (albacore, bigeye classified by size, but not by species) | <ul style="list-style-type: none"> - Biologists become involved in central statistical office so that persons responsible for statistics realise problems and correctly identify species reported by local names. - Train local statistical staff and fishermen to report using correct names |
| 4. Groups of species reported together. No column provided on reporting forms to report a certain species. | <ul style="list-style-type: none"> - Many countries where tuna fisheries are of minor importance, or where one species of tuna is minor to the other | <ul style="list-style-type: none"> - Add a column on reporting form to report species in question |
| <i>Inadequate coverage</i> | | |
| 1. Landings at some ports are missing. Survey system does not cover them. | <ul style="list-style-type: none"> - Many countries | <ul style="list-style-type: none"> - Expand survey system - Make occasional visits to those ports for landing estimates |
| 2. Certain fleets (most likely very localised) are missing. | <ul style="list-style-type: none"> - Many countries | <ul style="list-style-type: none"> - Raise catches to 100% coverage, using ratio of no. of boats covered to those not covered |
| 3. Landings at foreign ports are not covered. | <ul style="list-style-type: none"> - Panamanian fleets | <ul style="list-style-type: none"> - Oblige captains of boats to report catches by law |
| 4. Catch was used for family consumption or not sold through markets. | <ul style="list-style-type: none"> - Almost all countries | <ul style="list-style-type: none"> - Make sample survey to estimate amount |
| 5. Catch is sold on the local market. | | <ul style="list-style-type: none"> - Make survey by team of non-governmental personnel to make estimates |
| 6. Catch is classified and reported mixed with non-tuna species. | <ul style="list-style-type: none"> - Many countries in the Mediterranean, Africa, South America and Caribbean Sea | <ul style="list-style-type: none"> - Establish system (and format) to report tuna - Sample mixed species |
| 7. Catch is transhipped at sea from a fishing vessel to another fishing or cargo vessel of the same or different flag. | <ul style="list-style-type: none"> - U.S., Spain, Japan etc. | <ul style="list-style-type: none"> - Oblige captains to report catches regardless of landing or transhipping - Check landing ports (e.g. Puerto Rico) for cargo and foreign transhipments |

| | | |
|---|--|--|
| 8. Catch is landed at a customs zone and from there exported to a foreign country (often after canning). | - Many African ports - Foreign flag vessels unloading at Canary Islands | - Check landings at custom zones (do not use through-custom statistics) |
| 9. Catch is processed on the motherboat and is landed as canned product. | - Japan, former U.S.S.R | - Oblige captains to report catches instead of using landing statistics |
| Flags | | |
| 1. Double reporting by flag country, country that licenses the boats, importer of catch and/or country where fish are transhipped. | - Fleets of Korea, Panama, Japan, Ghana etc. landing at African ports | - People involved should realise the problem - Each country involved should report catches by flag - Secretariat monitors the flow of fish and fleets |
| 2. No reporting for boats where different nationalities of owners, operators, crew, investors, licensers, registration etc. are involved. | - Panamanian flagged vessels | - Government of operators of boats could encourage or request reporting of catch of such foreign flag boats (catch should be reported separately by flag to avoid confusion) |

4.2.6 Estimation and raising

Estimating for the observable population

Estimation of means, totals, and variances from sample surveys of logbooks, landings, or fishing trips by observers can be achieved with the standard estimation formulae appropriate to the sampling scheme used to draw the samples. Examples of statistical sampling schemes corresponding to those in the text books were listed in section 4.2.2. The estimated statistics then apply to the observable population (cf. section 4.2.1) without additional raising. Using inappropriate formulae, e.g. using simple random sampling formulae for a probability proportional to size sample, would risk biasing the results. Four examples are given to clarify these statements in a fishery context:

1. Suppose that n landings are sampled approximately randomly (simple random sampling without replacement) from N made by the fleet during a quarter of a year, and total numbers of fish, y , and length frequency distributions are estimated for each observed landing. The mean number per landing is $\bar{y} = \sum y/n$ and the estimated total landed number for the fleet is $Y = \bar{y}N$ (Thompson, 1992, eq. 8, chapter 2). The raising factor is thus N/n . It could be used to raise the LFD as well if required.
2. Suppose that n fishing trips were accompanied by observers during a year, and that the trips were selected without replacement from N made by the fleet so that each trip had an approximately equal chance of being observed. The raising factor is N/n .
3. As a contrast to 2), suppose an observer survey in which v vessels were selected randomly without replacement from V in the fleet, and no more than one randomly selected trip was observed on each. Firstly, the total quantity caught in the year by each observed vessel is estimated by raising the quantity observed on the trip by the annual number of trips by that vessel. Next, the total caught by the fleet is estimated by raising the vessel annual totals by V/v .
4. A logbook survey in which not all vessels could be logged could be treated as example 3). The annual results for v logged vessels should be available directly from the logs. They would be raised by V/v to allow for the missing vessels.

Estimators applicable to standard sampling designs without assumptions are arguably not the most precise way of estimating for the observable population when good auxiliary information is available. Auxiliary information

is brought into estimation by modelling using either a ratio estimator or a regression estimator (Thompson, 1992, chapters 7 and 8). Both are slightly design-biased, but model-unbiased.

For example 1), suppose that the total tonnage, W , landed by the fleet is known. It can be argued that the relative tonnages, w , of landings for the sample of n and for the whole fleet give a more precise raising factor than the numbers of landings because they give more information about the activities of the fleet. A ratio, rather than a regression estimator is appropriate if zero landed numbers always weigh zero tonnes, and positive landed numbers are proportional to tonnage. The ratio estimate of total numbers landed is $Y_{\text{ratio}} = W \cdot \sum y / \sum w$.

The raising factor is $W / \sum w$.

For example 2), a raising factor based on the quantities of fish landed by the fleet, and retained on the n observed trips may be thought more precise than one based on relative trip numbers. Raising by landings does however bring its own problems:

- If the species was not landed there is no raising factor unless composites of different species are used.
- If retained quantities were small on the observed trips the raising factor will be very imprecise due to the small divisor.
- If landed quantities are as weights and retained fish are as numbers, a conversion must be applied which could add further error.
- If landings are themselves suspected to be in error, the observer data become contaminated with the same errors.

Raising by relative fishing effort will often be preferable for observer trips if suitable data are available. The raised data are then independent of the landings data that can be important for modelling the fish stock using both types of data.

For example 3), relative effort might be used to replace total numbers of trips when estimating the annual total per vessel, whilst relative engine power might be used instead of V/v .

For the logbook example 4), a raising factor based on numbers of vessels logged and in the total fleet uses no information about the fishing powers of the different vessels and could often be improved. Other estimators could be based on engine power, GRT, days spent at sea or another effort measure, or on quantities of one or more species landed by the logged and by the total fleet. The choice will often depend on the information available and on whether the same information is being used in a linked way in an analysis of the fishery or stock.

Whatever estimation and raising formulae are used in preparing estimates, they should be documented in the SOP for the sampling programme.

Estimating for the population of interest

As pointed out in section 4.2.1, the observable population will probably be a subset of the population of interest in many practical situations associated with ocean fisheries and a relationship between the two has to be assumed (and documented in an SOP). The relationship may be a simple raising factor or something more elaborate but, in either case, estimation is purely a modelling exercise that does not involve design-based sampling theory. The models should be reviewed whenever new information becomes available.

A common but un-straightforward raising problem arises when the observable population for sampling is all or part of a fishing fleet but the population of interest is the stock of fish in a geographic region over a given time period, e.g. a 'time-area stratum'. Two questions arise:

1. Did the fleet fish the time-area stratum so as to provide a satisfactory sample of the fish stock there?
2. Was a part of the fishing effort expended outside the stratum?

Question 1) requires an analysis of fishing patterns and fishing gears in use. If the fleet distributed effort fairly uniformly over the stratum, and used similar fishing gear, a good sample was obtained. More likely though, parts of the fleet focussed on different, relatively small localities and used different gears having different selectivity

properties. There may also have been seasonal changes in fishing patterns. Options for dealing with these problems include:

- A model to estimate the effects of non-uniform sampling (Campbell, 2004). A Bayesian approach is worth considering because prior distributions for important parameters used in the model can be adjusted to reflect uncertainty about their values.
- Down-weighting results from heavily fished, small localities so that they do not dominate the estimates for the time-area stratum.
- Filling in missing data for sub-areas or sub-periods with data from previous periods or from neighbouring strata.
- Ignore all the irregularities and treat as a random sample.

All of the options carry a high risk of bias. Also, the bias could itself follow a trend over time resulting in distortion of time-series. A further danger is that data from certain useful locations will be used more than once causing dependence among results, and possibly exaggerating the effects of errors. The chosen solution and the reasons for choosing it should be documented in an SOP or in any report of the estimation procedure.

Question 2) is usually easier to deal with. The best solution is to disaggregate data for fishing trips into data for individual sets that can be assigned to the correct stratum. A quicker solution is to estimate for each trip the proportion of effort applied within each stratum.

Calculating the fraction of the total catch weight of a set represented by sampling

As noted in the section on sampling multi-species catches, the effective sampling unit assigned should be representative of the set. If a set is emptied into two or more wells of a purse seiner, the ratio that the weight in each well represents of the total weight caught in that set should be calculated. This is a weighted catch (Sarralde et al., 2005).

The sampled weighting of the catch cannot be calculated until the boat has finished unloading, as the catch in that set is spread over different wells and hence the individual set may have been sampled more than once.

The weighted catch of each set, defined by the date of the fishery and number of sets, can be calculated as:

$$\text{Weighting} = \frac{W1}{W2} * TW$$

where W1 is the weight of the set or sets in the well, W2 is the weight of the set or sets for all the wells sampled, and TW is the total weight of the set or sets.

For example, a vessel has caught 90 t in a set, which has been kept in three wells. 40 t were placed in well 1, 30 t in well 2, and 20 t in well 3. Wells 2 and 3 have been sampled.

For well 2, the weighting is: $(30/50)*90 = 54$ t.

For well 3, the weighting is: $(20/50)*90 = 36$ t.

$54+36=90$.

If the set had been sampled once only (either because the entire contents have been emptied into one well or shared over other wells that were not sampled), the weighted catch would be the same as the total catch of the set (weighting = TW). Samples can then be raised to the total catch of a set using this weighting.

4.2.7 Further reading

- BEARE, D., J. Castro, J. Cotter, O. van Keeken, L. Kell, A. Laurec, J.C. Mahé, S. Munich-Peterson, J.R. Nelson, G. Piet, J. Simmonds, D. Skagen and P.J. Sparre. 2002. Evaluation of research surveys in relation to management advice (EVARES). Final report of FISH/2001/02-Lot 1, DGXIV Fisheries, European Commission, Brussels. Available from a.j.cotter@cefias.co.uk.
- BOHRNSTEDT, G.W. and A.S. Goldberger. 1969. On the exact covariance of products of random variables. *Journal of the American Statistical Association* 64: 1439-1442.
- CAMPBELL, R.A. 2004. CPUE standardisation and the construction of indices of stock abundance in a spatially varying fishery using general linear models. *Fish. Res.* 70(2-3): 209-227.
- COCHRAN, W.G. 1977. *Sampling techniques*. New York, J. Wiley & Sons, Inc.
- COTTER, A.J.R., G.P. Course, S.T. Buckland, C. Garrod. 2002. APPS sample survey of English fishing vessels to estimate discarding and retention of North Sea cod, haddock, and whiting. *Fish. Res.* 55: 25-35.
- LUO, Z.-j, T. Tanaka, F. Kimura, M. Miyasaka, G.P. Course, A.J.R. Cotter, R.O. Jolliffe, B.J. Kay, M.S. Rolfe. 1999. Safe practices for sampling commercial trawler catches at sea. *Fish. Res.* 41: 99-104.
- DAVISON, A.C. and D.V. Hinkley. 1997. *Bootstrap methods and their application*. Cambridge University Press, Cambridge UK
- EFRON, B. and R.J. Tibshirani. 1993. *An introduction to the bootstrap*. London, Chapman & Hall
- GAERTNER, D., P. Pallarés, J. Ariz, A. Delgado de Molina, and V. Nordström-Fonteneau. 2000. Estimación de la durée des calées chez les senneurs français et espagnols opérant dans l'océan Atlantique, à partir des observations scientifiques du programme européen sur le patudo (1997-1999). *Col. Vol. Sci. Pap. ICCAT*, 51(1): 402-415
- GAERTNER, D., R. Pianet, J. Ariz, A. Delgado de Molina and P. Pallarés. 2003. Estimates of incidental catches of billfishes taken by the European tuna purse seine fishery in the Atlantic Ocean (1991-2000). *Col. Vol. Sci. Pap. ICCAT*, 55(2): 502-510.
- GOODMAN, L.A. 1960. On the exact variance of products. *Journal of the American Statistical Association* 55: 708-713.
- GOODMAN, L.A. 1962. The variance of the product of K random variables. *Journal of the American Statistical Association* 57: 54-60.
- GUTHRIE, D., J.M. Hoenig, M. Holliday, C.M. Jones, M.J. Mills, S.A. Moberly, K.H. Pollock, and D.R. Talhelm (eds.) 1991. *Creel and Angler Surveys in Fisheries Management*. American Fisheries Society Symposium 12.
- HILBORN, R. and C. Walters. 1992. *Quantitative Fisheries Stock Assessment-Choice, Dynamics and Uncertainty*, Kluwer Academic Publishers, Boston. 570.
- KENDALL, M.A. 1976. *Time-series*. London, Charles Griffin and Co. Ltd. 197.
- KIMURA, D.K. 1977. Statistical assessment of the age-length key. *J. Fish. Res. Board Can.* 34: 317-324.
- LAI, H.-L. 1993. Optimal sampling design for using the age-length key to estimate age composition of a fish population. *Fish. Bull.* 92: 382-388.
- PALLARÉS, P. and P. Dewals. 1998. Análisis de los datos obtenidos en una experiencia de muestreo intensivo de una cuba durante el desembarco. *Col. Vol. Sci. Pap. ICCAT*, 48(2): 213-220.
- PALLARÉS, P. and Ch. Petit. 1998. Tropical tunas: new sampling scheme and data processing strategy for estimating the composition of catches by species and sizes. *Col. Vol. Sci. Pap. ICCAT*, 48(2): 230-246.
- PALLARÉS, P., A. Delgado de Molina, J. Ariz and J.C. Santana. 2000. Revisión de las estadísticas españolas de túnidos tropicales (1991-1996), teniendo en cuenta el tipo de asociación de las pescas. *Col. Vol. Sci. Pap. ICCAT*, 51(1): 378-386.
- PINHEIRO, J. C. and D. M. Bates. 2000. *Mixed-effects models in S and S-plus*. New York, Springer. 528.

- RAJ, D. 1968. Sampling theory. New York, McGraw-Hill.
- ROSE, G.A. and D.W. Kulka. 1999. Hyperaggregation of fish and fisheries: how catch-per-unit-effort increased as the northern cod (*Gadus morhua*) declined. Can. J. Fish. Aquat. Sci. 56(Suppl. 1): 118-127.
- SARRALDE, R., A. Delgado de Molina, J. Ariz, J.C. Santana, P. Pallarés, R. Pianet, P. Dewals, A. Herve, R. Dedo and J.J. Areso. 2005. Port sampling procedures for tropical tuna in the Atlantic and Indian Oceans. SCRS/2005/101.
- SMITH, P.J. 1989. Is two-phase sampling really better for estimating age composition? Journal of the American Statistical Association 84: 916-921.
- SUKHATME, P.V. and B.V. Sukhatme. 1970. Sampling theory of surveys with applications. Ames, Iowa, USA, Iowa State University Press.
- SWAIN, D.P. and A.F. Sinclair. 1994. Fish distribution and catchability: what is the appropriate measure of distributions? Can. J. Fish. Aquat. Sci. 51: 1046-1054.
- THOMPSON, S. K. 1992. Sampling, John Wiley & Sons, Inc. 343.
- WESTRHEIM, S.J. and W.E. Ricker. 1978. Bias in using an age-length key to estimate age-frequency distributions. J. Fish. Res. Board Can. 35(2): 184-189.