QUALITY VERSUS QUANTITY: A CASE STUDY

J. Chesson
D.J. Staples

Fisheries Resources Branch
Bureau of Resource Sciences
PO Box E11 Queen Victoria Terrace
Parkes ACT 2600

What is data quality?

Data quality is not a simple concept. Quality can be measured directly by calculating the probability that a decision based on the data is 'correct'. In statistical terms this is referred to as the Type I and Type II, or false positive and false negative, error rates. While this measure is useful for well-defined situations, for example, determining compliance with a water discharge standard, less direct measures are necessary in the majority of cases. One suite of measures that is commonly applied is given in Table 1.

Clearly, any particular set of data can score highly with respect to some measures and poorly with respect to others. Also, some measures, such as bias and precision, can be applied to individual data values, while others, such as completeness, apply to the set of data as a whole. Some measures, such as comparability, can be applied to either. Often the quality of a statistic derived from a set of data (eg a mean or a correlation coefficient) is of primary interest. We will illustrate the application of these data quality measures to fisheries data using a recent redfish stock assessment as an example.

Example

Redfish (*Centroberyx affinis*) is one of the 16 species of the South East Fishery (SEF) managed by a total allowable catch (TAC) and individual transferable quotas (ITQs). Redfish is a convenient example because it was the subject of a stock assessment workshop earlier in 1993 (SEFSAG 1993). The choice of redfish is not meant to suggest that the redfish data are particularly problematical. In fact the existence of data over such a long time period is a credit to the institutions and individuals involved. We expect that the points illustrated by redfish are equally applicable to many other species.

Tables 2, 3 and 4 provide a qualitative assessment of how well three types of redfish data score against each of the five data quality measures in Table 1. The first step in such an assessment is to determine the purpose for which the data are to be used. The same set of data may score well for one purpose and poorly for another.

Total catch

Table 2 evaluates landings data in the context of determining the total catch of redfish. Total catch is needed for evaluating mortality due to fishing and calculating important indicators such as catch per unit effort. The most significant data quality issue here is the, possibly large, negative bias introduced during periods when there was virtually no market for redfish and a significant portion of the catch was discarded at sea (Figure 1). The impact of such a bias is shown in Figure 2. When landings data are taken at face value, the biomasses estimated by the virtual population analysis (VPA) program CATANAL (Walters 1987) follow the 'no discard' curve. When the data are adjusted to include a hypothetical rate of discarding set at 80 percent for 1969 to 1980 and then declining to 10 percent by 1987, the 'high discard' curve is obtained. The estimated biomasses are quite different, particularly for earlier years. Depending on which scenario is closer to the truth, one gets a different impression of the degree to which the stock has been depleted and the potential biomass if the stock were able to return to earlier levels. Unfortunately there is no quantitative record of the actual level of discarding. Without such a record, the quality of the landings data for the purpose of estimating catch is low.

Length frequency

Table 3 evaluates length frequencies recorded at the Sydney Fish Market with respect to determining the length frequency of the catch. A key issue here is representativeness. Although the data are likely to be relatively unbiased with respect to fish reaching the market, and of reasonably high precision, they may not be a representative sample of the entire catch which has both retained and discarded components. Data from the recently established Scientific Monitoring Program (see below) show that the length frequency distributions for the two components can be quite different (Figure 3). Using the distribution from the retained component as representative of the entire catch could be quite misleading.

Another dimension of data quality emerges when we revise our purpose from determining the length frequency of the catch to determining the length frequency of the stock itself. Selective fishing at particular locations and depths can mean that the catch is not representative of the stock. Fishers have claimed that the decline in mean length of redfish reaching market is a result of fishing at shallower depths rather than a change in population structure. Limited data on length frequency by depth from FRV Kapala surveys combined with logbook records enabled us to predict the trend in mean length if change in depth were the only factor. Comparison of the predicted trend with the observed trend (Figure 4) shows that at least some (but not all) of the apparent changes in mean length can be explained by changes in the depth of fishing. Data that may be of the highest quality with respect to determining the length frequency of the catch, may be less suitable for determining the length frequency of the stock. Unless changes in fishing practice such as fishing at shallower depths are taken into account, changes in the length frequency of the catch may be incorrectly attributed to changes in the length frequency distribution of the stock.

Catch per unit effort (cpue)

Table 4 evaluates the quality of cpue data with respect to obtaining an index of stock abundance. The evaluation is quite superficial and does not consider the relationship between cpue and abundance. In other words, the table concentrates more on the quality of cpue data for determining cpue rather than abundance. (Cpue can be of interest in its own right for economic analyses or studies of fisher behaviour.) Comparability is an important data quality issue here. Changes in the types of boats and fishing practices over the years make it difficult to compare different measures of effort.

If cpue is to be used to provide information on abundance, the temporal and spatial distribution of effort must also be taken into account. For example, if cpue increases from one year to the next, but fishers have moved to a new fishing ground, the change in cpue may reflect the spatial distribution of fish rather than an increase in their abundance.

Messages

The main messages from the redfish example can be summarised as follows:

- Data quality has to be determined within the context of a particular use. The issue is whether or not data are of sufficient quality for an intended use, not simply whether the data are good or bad..
- Data quality is multi-dimensional. A data set may score well on some measures and poorly on others.
- Data quality can change (usually deteriorate) over time. Important ancillary information that can help determine, and in some cases correct, problems of data quality is easily lost if data are not properly stored and documented.

Improving data quality does not necessarily involve additional input of resources. It requires:

- planning for optimum resource allocation (taking into account the strengths and weaknesses of different types of information).
- regular monitoring and assessment of data quality.
- documentation of data and measures of its quality care of data so that all of the above are not lost.

Actions

In conclusion, we will describe two actions currently under way that are aimed at improving data quality for fisheries resource assessment.

The first is the South East Fishery Scientific Monitoring Program. The program which is funded by the Fisheries Resources Research Fund and coordinated by the Bureau of Resource Sciences in cooperation with New South Wales, Victoria and Tasmania, places scientific observers on domestic fishing vessels. The observers collect information on both the retained and discarded portions of the catch. Table 5 indicates some of the strengths and weaknesses of the Program in terms of data quality. In order to save space, we have violated our earlier requirement that data quality should be considered in the context of a specific objective, and have used a generic 'everything' to refer to various types of fisheries information.

The Program collects data from a sample of fishing trips, not from every boat or every trip. Precision is reduced because of the additional variability introduced through sampling rather than obtaining data from a complete census in which all trips are included. On the other hand, the Program enables us to collect information such as the quantity and type of discarding that is not available by any other means. Figure 3 is one example. By proper random or stratified random sampling, some sources of bias can be controlled. The size of the sample (number of trips) can be adjusted to achieve a balance between the precision required and the resources available. Representativeness is an issue because the Program only operates out of certain ports which may or may not be representative of the fishery as a whole. Continuing resources and a high standard of documentation and data management are necessary to satisfy completeness and comparability requirements.

The second action, which is closely linked to the Scientific Monitoring Program, is the formation of the South East Fishery Stock Assessment Group and its plans to review data quality needs. Many of the quantities used for stock assessment can be obtained in more than one way. Figure 5 shows a variety of ways of obtaining total catch. Each way has advantages and disadvantages in terms of data quality and they are likely to vary from species to species and between fisheries. The Stock Assessment Group intends to investigate the relative merits of different approaches, including the advantages of using more than one approach in some cases, in order to make the best allocation of resources across different data collection activities. Obviously, it is of little benefit to spend a lot of time and effort improving the quality of one type of data and ignoring the quality of some other piece of information with which it has to be combined.

The aim of exercises such as this is to ensure that the quality of data is appropriate for the task at hand. If data quality is too low, the ability to make decisions based on those data is compromised. If some aspect of data quality is unnecessarily high, it probably means that resources have been poorly allocated. Clearly it is easier to address data quality for specific projects. When data are being collected for a number of possible uses, it is more difficult to specify the ideal mix of data quality attributes. Nevertheless, it is important that a formal consideration of data quality be an integral part of any data collection activity. Otherwise, it is likely that resources will be wasted collecting data of inappropriate quality.

References

SEFSAG (1993). Redfish (*Centroberyx affinis*) Stock Assessment Report 1993. Draft August 1993. Compiled by Bureau of Resource Sciences for the South East Fishery Stock Assessment Group, Canberra.

Walters, C.J. (1987). *Microcomputer catch-at-age analysis system*. Resource Ecology, University of British Columbia, Vancouver.

Table 1. Data quality measures

- bias: the tendency to be bigger (smaller) than the real thing;
- precision: how close repeated measurements are to one another;
- representativeness: the degree to which the data represent the population of interest;
- completeness: absence of gaps or missing information;
- comparability: ability to compare with other data.

Note: "Accuracy" is often used to mean zero bias. In other situations it is used to refer to a combination of small bias and high precision. Because of this ambiguity, we avoid using "accuracy" in this paper.

Table 2. Assessment of data quality with respect to total redfish catch

What we want to know: **total catch.**

What we have: **landings data from 1937 to 1991.**

bias	severe negative bias during some periods. Redfish discarded at sea. Fish sold on black market.
precision	Unknown, but probably reasonable if based on market weighings.
representativeness	Probably quite good. Expected to include most of the redfish fishery.
completeness	time series for 1937 to 1991.
comparability	More recent data are from log books rather than landings. May suffer from different data quality problems (eg. underreporting).

Table 3. Assessment of data quality with respect to length frequency of redfish catch

What we want to know: **length frequency of catch.**

What we have: **length frequencies from Sydney Fish Market, 1947–1991.**

bias	Unknown, but probably small. Assume measurements are a random sample of fish reaching market.
precision	Probably quite good. Measurement process straightforward. Apparently done by same individual for many years.
representativeness	Market fish not representative of catch during periods of selective discarding. Also, catch is not representative of redfish stock when fishers target particular locations, depths, etc.
completeness	Gaps in time series, essentially no data for 1968 through 1973. Measurements discontinued after 1991.
comparability	changes in measurement method (eg. from reporting nearest cm to reporting nearest cm below) must be documented and taken into account when comparing data.

Table 4. Assessment of data quality with respect to abundance of redfish stock

What we want to know: **index of stock abundance.**

What we have: **various cpue series covering 1918 to 1991.**

bias	Unknown.
Precision	Unknown.
representativeness	Probably quite good with respect to cpue itself. Most boats included. Relationship between cpue and abundance has to be explored.
completeness	Multiple series, with only limited overlaps.
comparability	Significant difficulties in comparing different measures used at different times for different types of boat, fishing methods, etc.

Table 5. Assessment of data quality issues associated with the scientific monitoring program

What we want to know: **everything.**

What we have: **Scientific Monitoring Program.**

bias	Small for ports being monitored, provided samples are selected according to a proper sampling scheme (eg. random or stratified random sampling). Able to record both retained and discarded portions of catch. May be significant sources of bias for fishery-wide quantities if sample is not sufficiently representative of entire fishery (see below).
precision	Sampling variability reduces precision compared to census approach. Magnitude depends on variability of quantity being measured and number of samples. Under evaluation.
representativeness	Only covers limited number of ports. May not represent entire SEF.
completeness	Needs continuing resources to ensure high degree of completeness.
comparability	High standard of documentation and data management required to maximise comparability with other information.

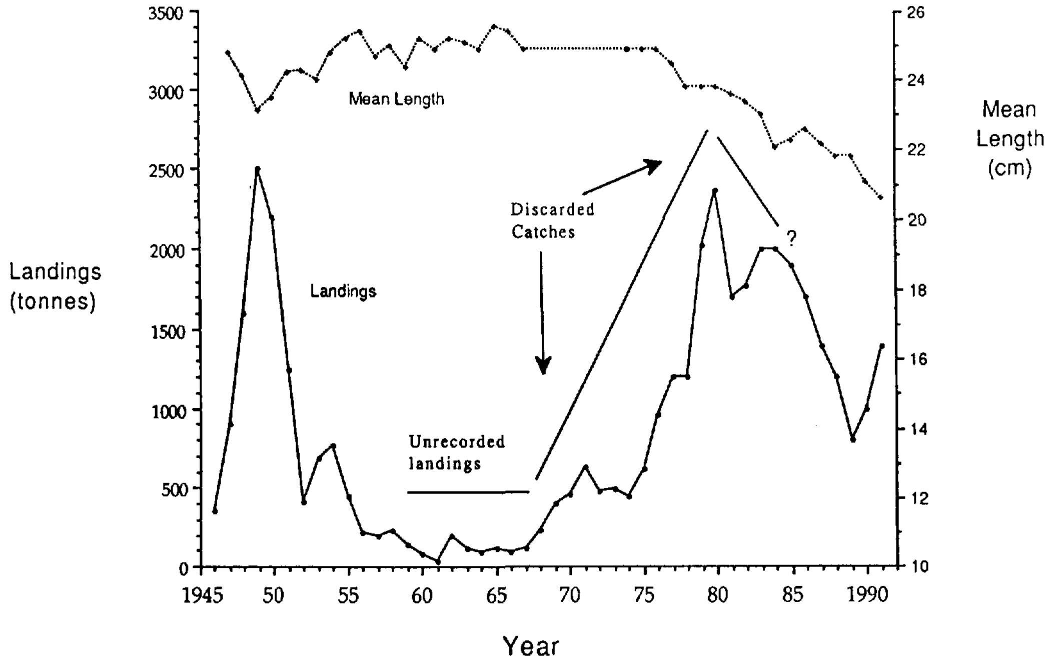

Figure 1. Annual landings of redfish and mean length of redfish marketed at the Sydney Fish Market, 1946 - 1991. (Source: Kevin Rowling, NSW Fisheries Research Institute)

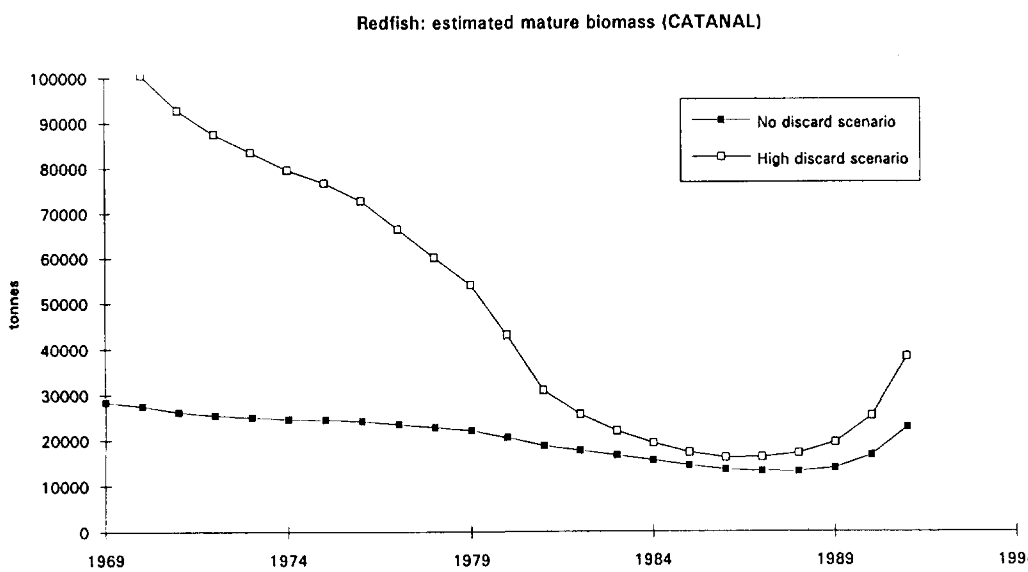

Figure 2. The effect of discarding on estimates of redfish mature biomass obtained from the virtual population analysis program CATANAL.

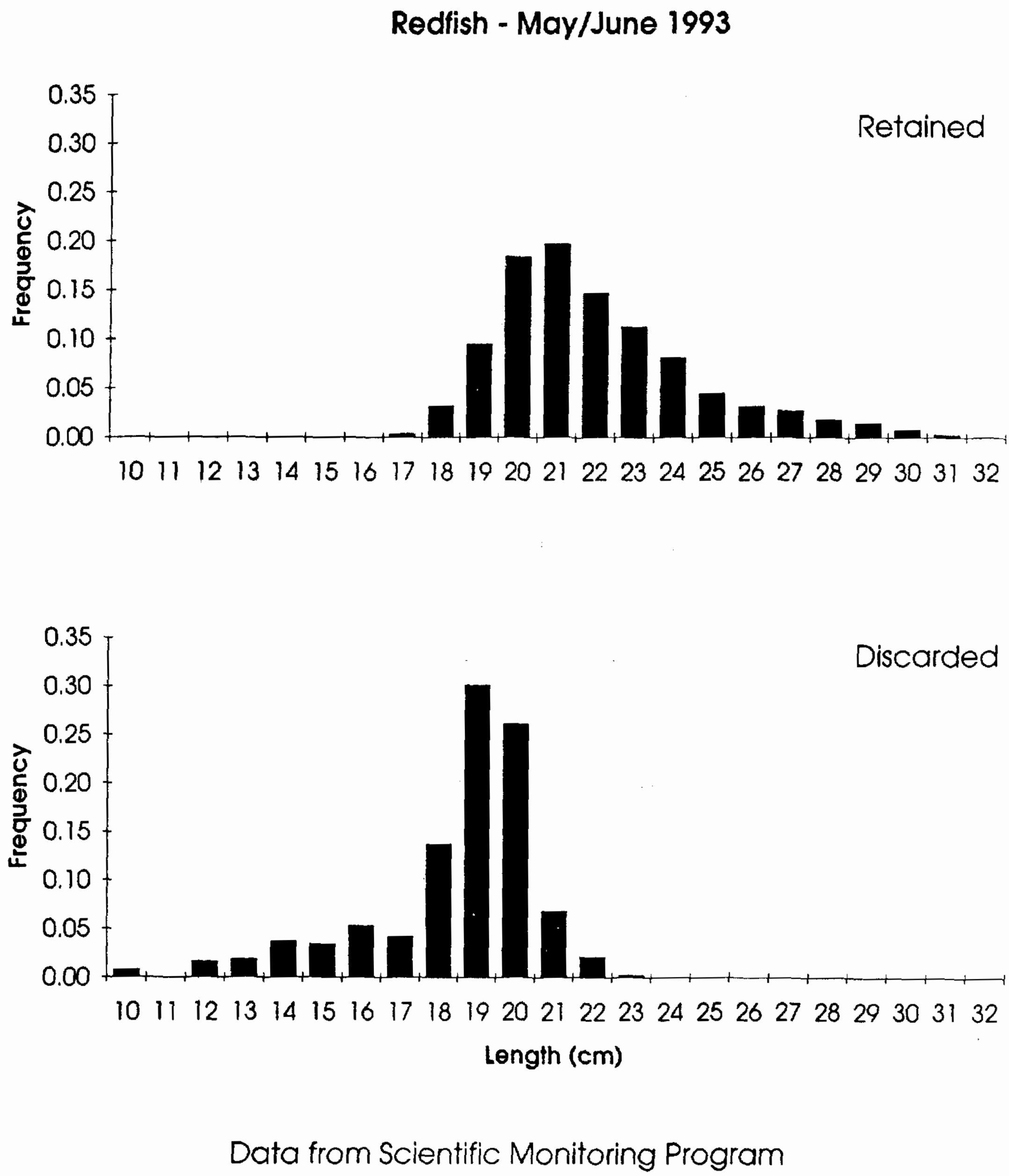

Figure 3. Distribution of redfish lengths in the retained and discarded portions of the catch. Data from the Scientific Monitoring Program.

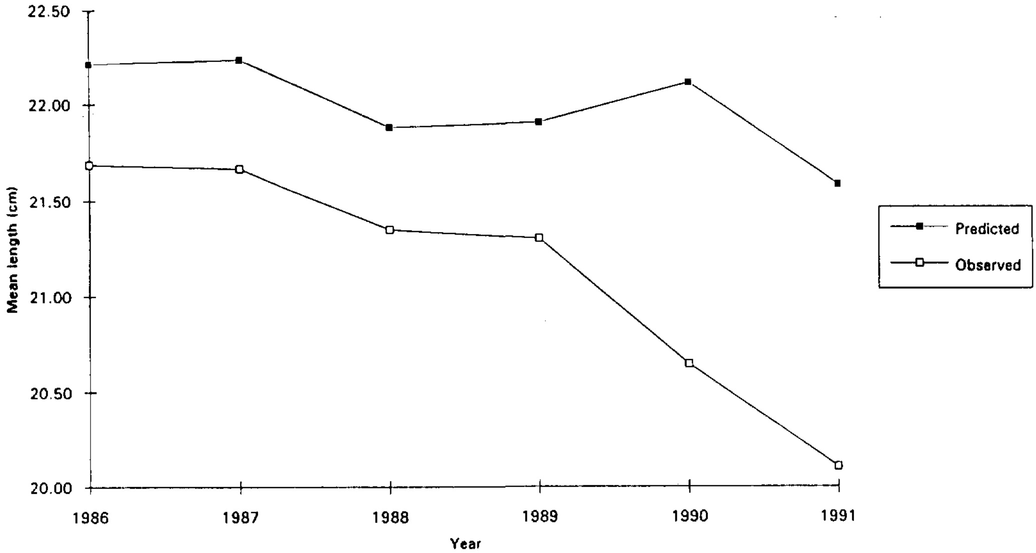

Figure 4. Mean length of redfish predicted from fishing depth compared with observed mean length.

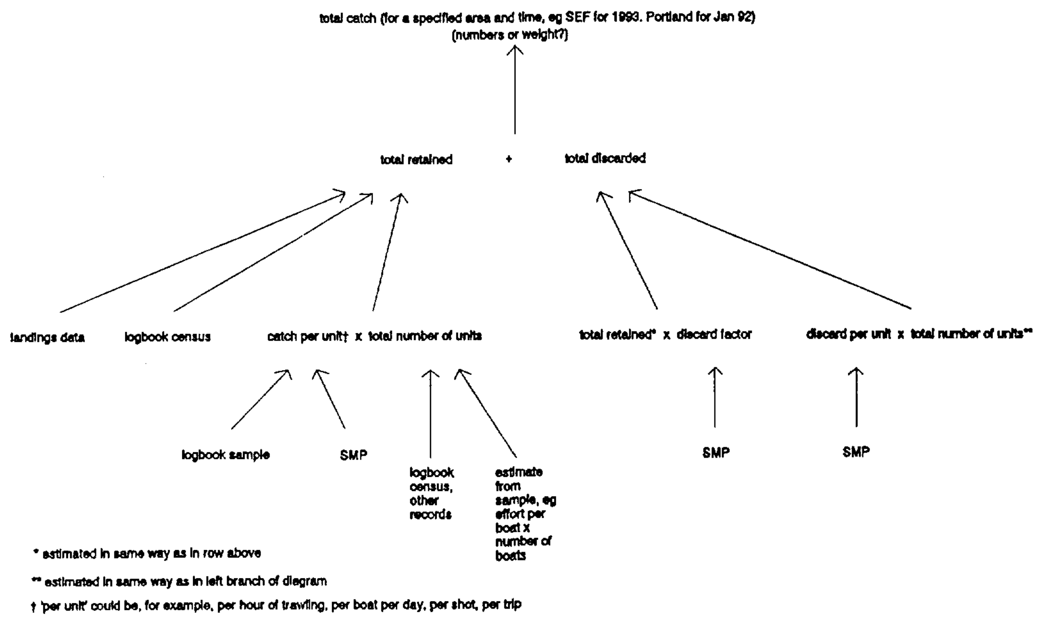

Figure 5. Some of the possible options for obtaining an estimate of total catch.