CHAPTER ONE: INTRODUCTION

Melita Samoilys and Neil Gribble

1.1 Background and aims

The primary aim of this manual is to enable fisheries officers and other interested scientists to monitor fish stocks that are exploited on Pacific coral reefs, and thereby make some predictions on the status of those stocks. The manual addresses the practical aspects of monitoring exploited coral reef fish stocks. In developing a comprehensive monitoring technique, fisheries scientists and other researchers will be able to collect data that is vital to fish stock assessment. This manual focusses on providing procedures for collecting reliable data and guidelines for interpreting such data. The latter falls under the category of stock assessment, the primary goal of fisheries science.

The following section (1.2) provides an introductory overview of fisheries stock assessment. This field is constantly evolving, and at present lacks a strong consensus on appropriate models for assessing the complex multi-species, multi-gear fisheries typical of coral reefs. Basic stock assessment approaches are based on single-species surplus production models (SPM) or yield per recruit models (YPR), both of which are generally considered too simplistic for coral reef fisheries. Alternatively, multi-species models which account for species interactions such as predator-prey relationships, are complex and require a prohibitive amount of data (Appeldoorn 1996), which are typically unavailable in countries where coral reef fisheries occur. These fisheries are difficult to assess because they are multi-species, multi-trophic and are characterised by a wide range of fishing methods and multiple landing stations.

Within this framework we aim to describe techniques for obtaining reliable estimates of basic parameters needed to describe a tropical multi-species finfish fishery such as stock abundance, catch and effort, on the basis that such data can be used *ad infinitum* as new models are developed and existing models evolve. At present, some useful stock assessment procedures have been developed which involve a combination of small-scale lumping (combining or grouping samples) and simple single species models (Appeldoorn 1996, Polunin et al 1996). These procedures can be applied to the types of data collected by the methods described in this manual to provide useful management information.

This manual focusses on underwater visual census (UVC) surveys of stock abundance which are fishery-independent methods, and catch per unit effort (CPUE) surveys to obtain catch and effort data, which are fishery-dependent methods. Estimates of stock abundance and CPUE are used to detect trends or perturbations in stocks. Such estimates may also be used to predict potential yield and the health of stocks. The manual's emphasis is on accurate and rigorous methodology in the collection, storage, management, analysis, interpretation and presentation of data. A major aim of the manual is to provide fisheries officers with methods for collecting reliable and consistent stock abundance, catch and effort data over time, so that they can accumulate a time-series of data to monitor their coral reef fisheries. Readers should refer to the recently published book "Reef Fisheries" edited by Polunin and Roberts (1996) for a detailed and thorough synthesis of the current state of knowledge on coral reef fisheries.

The manual builds on a Queensland Department of Primary Industries (DPI) research project funded by the Australian Centre for International Agricultural Research (ACIAR) hereafter called the ACIAR/DPI UVC project, which investigated the suitability of UVC methods for fisheries stock assessment purposes. The ACIAR/DPI UVC project was a collaborative research project between Fisheries (DPI) in Queensland and the Fisheries Divisions of Fiji, Solomon Islands and Papua New Guinea, and is reported in Samoilys and Carlos (1992) and Samoilys et al (1995).

1.2 Fisheries stock assessment

Fish stock assessment at its simplest seeks to answer two basic questions:

What is the size of the stock (or how many fish are there)?

What is the *sustainable yield* from the stock (or how many fish can be caught while leaving enough to breed and build up numbers again)?

All stock assessment revolves around understanding and predicting parameters of *stock size* and *yield*.

A working definition of a fish *stock* is a population of a fish species where individuals have similar recruitment, growth and natural mortality (death rate) characteristics, and are genetically contiguous. These factors have a large

influence on productivity of a stock or fishery - hence if two populations differ significantly in these characteristics, which is invariably the case in a multi-species reef fishery (Appeldoorn 1996), they should be managed separately to ensure stock safety (Haddon and Willis 1995). The spatial definition of a stock must be large enough to incorporate movement of the fish, i.e. emigration and immigration are said to be negligible. The results of separate stock assessments may subsequently be pooled into an assessment of a multi-species fishery (see Sparre and Venema 1992). Identification of stocks in multi-species communities on tropical reefs can be a problem, but if disregarded there is the danger of unknowingly fishing down one stock while maintaining good catch rates over the combined stocks.

1.2.1 Abundance (stock size) measurements and indices

As noted above there are *fishery - dependent* and *fishery - independent* methods of estimating the size of a fish stock. Both sets of methods have inherent strengths and drawbacks. A combination of both will give the most reliable assessment of a fishery. A brief description of the basic methods of estimating fish abundance is outlined below.

Logbook CPUE (fishery - dependent data)

This method involves producing a spatial map of reported catch per unit effort (CPUE) in the fishery. CPUE, or the number of fish caught per day per boat, is assumed to be directly related to the abundance of fish. The next step is to calculate (as an extrapolation) the overall abundance of the target fish by averaging and summing the estimates of abundance in each local area. This gives a first approximation of the size of the stock for the total area of the fishery. It assumes the average CPUEs for the areas fished apply to the total area of the fishery. The drawback with logbook data is that commercial fishers target areas of high abundance hence calculations based on logbook data may overestimate stock size. This is particularly so if the target fish aggregate in schools or at spawning sites. In addition, reliability of logbook data in terms of truthful reporting, is unknown until the data have been validated (e.g. through observers on board commercial vessels).

Depletion or catchability studies (usually fishery - dependent data)

The catchability coefficient (q) is a measure of the ability of a given gear type to catch the target species present. Methods of calculating "q" include the Leslie or DeLury methods (Ricker 1975). The first involves plotting the CPUE against the cumulative catch over a period of time; the intercept gives the initial population or stock size and the slope gives the "q". The second method plots the log CPUE against the cumulative effort and the fitted straight line gives the same parameter estimates (Ricker 1975). Appeldoorn (1996) cites recent applications of these techniques to coral reef fisheries, and Samoilys et al (1995, chapter 9) report on depletion experiments in Fiji and Solomon Islands. An assumption made is that a population or stock can be fished until the CPUE drops, because CPUE is directly related to the abundance of the stock. This is not always the case, particularly when schools of fish are targeted. Here the CPUE will remain stable until the last fish in the school is caught and then there will be a dramatic drop in CPUE. Therefore the drawbacks to depletion methods are similar to those mentioned above: commercial fishers invariably target areas of high abundance - hence calculations may overestimate stock size. Again this is particularly so if the target fish aggregate in schools or at spawning sites.

Research surveys (fishery - independent data)

A spatial map of the distribution of the biomass or abundance of the target species is produced from the results of research trawling, line fishing or underwater visual census. Again it is possible to calculate (as an extrapolation) the overall abundance of the target fish from estimates of abundance in each local area, from which the average number per unit area is calculated and then extrapolated for the total fishery area. The drawback with this approach is that usually only relatively small areas can be surveyed adequately due to cost and time, which leads to uncertainty if results are extrapolated across a large fishery.

Tagging studies (both fishery - dependent and fishery - independent data)

These are forms of the dilution method of population estimation used in ecology (Ricker 1975). A small number of fish are tagged with visible markers and released. The ratio of marked to unmarked fish in subsequent catches gives an estimate of the ratio of the number of fish originally marked to the total abundance (e.g. Recksiek et al 1991). Drawbacks to this approach are the assumptions of complete mixing of marked fish within the whole population and that there is an equal probability of recapture. Both are unlikely with reef fish because of their non-random distribution and limited movement (Appeldoorn 1996, Samoilys 1997).

1.2.2 Population dynamics

The dynamics of a single species population can be simplified to the interaction of three factors: *recruitment* combined with *growth*, balanced by *mortality*. (It will be assumed that immigration and emigration either do not occur or have relatively minor effects, i.e. the dynamics of the population is investigated at the scale of a *stock*, see above). As recruits (into the fishery) grow, the combined biomass of the stock increases rapidly, usually much faster than the depletion due to mortality. When recruits reach adult size, growth slows and depletion of the population due to increased age-related mortality begins to decrease the biomass of the stock. The productivity of the fishery will be maximised if fishing occurs at or just before this point.

The following section gives brief descriptions of a variety of methods suitable for estimating parameters of population dynamics in single species systems.

Recruitment

In many exploited fish species, recruitment is the most variable element of productivity and therefore strongly influences the resilience of those populations to harvesting. For the purposes of stock assessment, recruitment is usually defined as: "***the number of juvenile fish that have attained the age (or size) when they become vulnerable to fishing gear***" (Sparre and Venema 1992). The timing and strength of recruitment can be determined by age/size frequency analysis of a time series from either commercial or survey samples. In the best case a large number of small fish (juvenile recruits) will be caught at only one time of the year (a year class or *cohort*). However, studies of larval settlement indicate some species recruit (into the population) continuously throughout the year or for substantial portions of the year (Doherty 1991), therefore estimation of recruitment can be difficult. Seasonality has been observed in the spawning and larval settlement of coral reef fishes at most geographic locations (Doherty and Williams 1988); typically larval settlement is restricted to fewer than five months over summer (Doherty 1991). Direct estimates can be made of larval settlement at the end of the reproductive season for tropical fish species (Doherty and Williams 1988). A strong correlation between survey counts of settling larvae and subsequent abundance has been identified for some species (Doherty and Fowler 1994).

Growth

The von Bertalanffy growth equation is the most commonly accepted function describing growth in commercially exploited marine animals.

$$L(t)=L_\infty \ [1\text{-}exp(\text{-}K*(t\text{-}t_0))]$$

where L_∞ is length at infinity (very old animal)

     K   is the slope constant (rate of growth)

     t_0 is time of zero length (initial condition parameter)

There are a number of variations on this theme usually involving an increasing number of parameters (e.g. Schnute 1981). A useful variation is the seasonally adjusted growth equation (Pauly and Gaschutz 1979, Somers 1988).

Mark and recapture method of growth estimation

Size at release is related to the size at recapture and the time at liberty. The method requires a good spread of "times at liberty" and sizes. The growth data is usually fitted to a von Bertalanffy equation via non-linear regression (e.g. Fabens (1965) algorithm).

Age based methods of growth estimation

Age readings are made from otolith rings, vertebrae cross-sections, spines, or scales. The length-at-age can then be tabulated and growth curves plotted (Sparre and Venema 1992 p51). Readers should refer to the extensive literature on ageing (e.g. Panella 1971, Beamish and McFarlane 1983, Francis et al 1992).

Length based methods of growth estimation

The length frequency time-series from a population can be used to derive growth data if there is no age data available. The average length of animals in a pseudo-cohort (distinct length class) can be followed through sequential samples, thus giving length-at-elapsed-time. The growth curves are therefore based on size rather than age classes. To a large extent these and the older graphical methods have been replaced by computer based modal/growth identification systems such as FISAT, ELEFAN, MULTIFAN or LFSA (see FAO/ICLARM publications and software). Their drawbacks relate to the assumption that modal size classes reflect cohorts.

Introduction

Mortality

Total mortality (Z) is made up of fishing mortality (F) plus natural mortality (M) ; i.e. Z = F + M. The instantaneous rate of mortality (i.e. natural log of the survival rate) is described by

$$N_t \Big/ N_0 = e^{-Zt}$$

or $\qquad Z = log_e N_0 - log_e N_t$

where N is the number of individuals before and after a given time interval "t" (Ricker 1975).

There are a number of methods for estimating Z from a time-series of research surveys or a combination of research and fishery landings data (see Sparre and Venema 1992, Appeldoorn 1996). One of the simplest is **Catch Curve** analysis. A graph is plotted of the logarithm of the number of the target species taken at successive ages (cohort) or sizes (pseudo-cohort). The latter giving length-converted catch curves (LCCCs, see Appeldoorn 1996). These data are derived from fishers' catch/effort logbook data, usually backed up with catch sampling or research surveys to establish size or age structure of the catch. Length data is converted to age data by the von Bertalanffy growth equation. The primary assumption is that the population is in equilibrium with respect to fishing pressure, i.e. there will be a rapid adjustment in the age structure of the stock in line with the rate of fishing. This adjustment will be reflected in the shape of the catch curve and therefore the slope of the linearised curve. Natural mortality is assumed to be constant through time and across all age classes. The slope of the linearised curve gives Z directly (Sparre and Venema 1992; see also Cumulated Catch Curve or the Jones and van Zalinge method, *op. cit.*).

Fishing mortality (F)

The instantaneous rate of fishing mortality is the ratio of fishing deaths to all deaths, multiplied by the instantaneous total mortality rate (Ricker 1975). The basic assumption is that fishing mortality relates directly to catch or CPUE and can be estimated from fleet catch/effort statistics. The relationship is F = fq where f is the fishing effort and q is the catchability coefficient. Estimation of "q" can be via depletion methods (described above).

Natural mortality (M)

M is usually calculated by simple manipulation of Z = F+M given that Z and F have been previously calculated. Tagging experiments during seasonal closures compared to similar experiments during the fishing season can be used to give independent estimates of M (Ricker 1975).

There are also a number of theoretical and empirical models that relate natural mortality to fish growth and age; for example Pauly's empirical formula (Pauly 1980). This assumes that a natural relationship exists between the rate of growth (K), the largest size attained (L¥), and the average sea-surface temperature, which will give the expected natural mortality (M) for a given target species. The original formulation was based on the regression of data on 175 different fish stocks (Pauly 1980).

1.2.3 Estimation of yield

The concept of **sustainable yield** is linked to that of **surplus production** from a fish stock. Surplus production is the proportion of the fish stock above that required for breeding maintenance of fish numbers. For example, in some species as the number of individuals in an area is reduced the breeding success of the remaining population may increase through density dependent population regulation. In such species the surplus population is available for harvesting without long-term detriment to the stock. A second example would be the taking of large fish after they have spawned at least once (e.g. through minimum size regulation in the fishery). Here the reproductive contribution of the animal has already been made and removing it reduces competition for resources with the next generation of juveniles. Again, theoretically, the harvest of these "surplus" individuals will not cause long-term detriment to the stock. However, these scenarios may be unrealistic for coral reef fishes and it may be difficult to identify the surplus component of the population. This is because many species are hermaphrodites, there is little evidence for density-dependent population regulation, and their population dynamics reflect highly variable recruitment and complex species interactions (see Sale 1991). These processes remain poorly understood for reef fishes, particularly the larger species typically exploited by fisheries. Nevertheless, in view of the present unavailability of alternative models, the concept of sustainable yield remains useful in providing a first order, though often over-optimistic, assessment of a reef fishery (see Chapter 6).

Production models (SPM)

The usual method of calculating yield or variants such as maximum sustainable yield (MSY) has been through

application of fairly robust models which incorporate a time-series of catch and effort statistics. There has been a general trend towards more sophisticated and complex stock assessment models as the quality and quantity of available data increases. However the simplest surplus production model (sometimes called the Schaefer model) uses the change in catch or yield per unit of effort with cumulative fishing effort to estimate the MSY; i.e. at some effort level the optimum yield-per-unit-effort will occur. This assumes that the relationship of yield to cumulative effort conforms to a simple theoretical curve function known as the Schaefer curve (see Sparre and Venema 1992, Appeldoorn 1996).

An underlying assumption of the traditional form of production models is that the stock is in *equilibrium*, where fish numbers are basically stable with increases due to recruitment and growth balanced by decreases due to a combination of natural and fishing mortality. Coral reef fish stocks are unlikely to be in equilibrium because their larval recruitment is highly variable (Doherty 1991). More recent innovations are the *biomass-dynamic* models which use maximum-likelihood techniques to estimate (or simulate) non-equilibrium situations (Hilborn and Walters 1992).

Approximate yield models

Gulland (1971) proposed a formula for estimating MSY by relating yield to the virgin biomass, assuming that fishing mortality at MSY is roughly equal to the natural mortality (see chapter 6). Garcia et al (1989) generalised the concept by taking into account the average exploited biomass rather than the virgin biomass, such that:

$$MSY = \frac{\overline{B}M^2}{2M - F}$$

Where $\overline{B}$ is the average exploited biomass

M is the natural mortality

F is the fishing mortality

Given the difficulty in assessing the fisheries potential of poorly documented reef-fish stocks, the use of the Gulland or Garcia et al models is recommended (Appeldoorn 1996). Estimates of biomass can be obtained via fisheries - independent research surveys such as UVC surveys.

Yield-per-recruit models (YPR)

These are a sub-set of the "dynamic pool" models which utilise the parameters of fish population dynamics rather than catch statistics. The YPR model follows a cohort of recruits through their life-history as they grow and die, until the fish are ultimately caught by the fishery. The ratio of the yield, as weight of fish caught, to the number of original recruits gives the YPR estimate. These calculations account for growth and mortality but not recruitment, therefore an optimum YPR may not be sustainable.

The general drawback to this family of models is that the predictions they give are only as good as the original assumptions and estimates of the population parameters. Unless care is taken at the parameter estimation stage the result can be a case of "garbage-in-garbage-out". Computer intensive techniques of parameter estimation such as linear programming, boot-strapping, and the use of Bayesian estimators are now used to refine stock assessment (see Hilborn and Walters 1992). However, the underlying biological relationships in the models remain the same.

Sustainability indicators

Rather than monitoring yield, stocks can be monitored via *sustainability indicators* such as *spawning biomass* or *recruitment strength*. Spawning biomass (or spawning stock biomass) is calculated as the number of fish alive multiplied by the fraction that is reproductively mature, in each age class, multiplied by the weight of an individual (Caddy and Mahon 1995; Laane and Peters 1993). Empirical and theoretical studies suggest that a stock-recruitment failure may occur when the spawning biomass of a finfish stock is fished to below 20% (Goodyear 1989, Plan Development Team 1990) of the unfished or *virgin spawning biomass*. A recent study gives a more conservative estimate of 30-40% (Caddy and Mahon 1995).

Recruitment strength has been addressed in the section on parameter estimates for population dynamics. Monitoring a time-series of such estimates allows early detection of a drop in recruitment, relative to previous years (see Caddy and Mahon 1995). If such a trend continues it indicates a potential stock-recruitment failure.

1.2.4   Summary and further reading

This introductory chapter provides a general overview of the process of fisheries stock assessment to introduce the reader to some of the main concepts. The manual does not address the application of these models, it addresses the estimation of certain parameters that stock assessment models require. More detail on any aspect discussed in this chapter can be gained from the following readily obtainable texts:

Introduction to tropical fish stock assessment
by P. Sparre and S.C. Venema (1992). FAO Fisheries Technical Paper 306/1
Quantitative Fisheries Stock Assessment: Choice, Dynamics & Uncertainty by R. Hilborn and C.J. Walters (1992). Chapman and Hall, New York
Adaptive Management of Renewable Resources
by C. Walters (1986). Macmillan, New York
Model and Method in Reef Fishery Assessment
by R. Appeldoorn (1996). In: **Reef Fisheries,** by N.V.C. Polunin and C.M. Roberts (eds) Chapman and Hall, London

1.3   How to use this manual

This manual details procedures for quantifying fish stocks exploited on coral reefs. The chapters are arranged in a logical order, with each chapter representing a key element in the chronological process of assessing exploited reef fish stocks. Chapter 2 discusses hypothesis testing and sampling design in research surveys; Chapters 3 and 4 describe the field-based sampling methods, UVC and CPUE surveys, respectively; Chapter 5 describes how to set up and manage a database and process data; Chapter 6 covers analysis and interpretation of data, and Chapter 7 defines the principles of reporting and presenting research survey results.

For most chapters the layout is designed to provide a hands-on field and desk guide. Set procedures are blocked and highlighted and these are followed by explanatory text which provides background information and references to relevant literature, so that the reader can explore the methods described in more detail. Important terms, definitions and key points are in bold and italic. At the back of the manual a field trip equipment checklist has been provided as a guide, with plenty of space for additions and further notes.

CHAPTER TWO:
SAMPLING DESIGN AND HYPOTHESIS TESTING

Marcus Lincoln Smith and Melita Samoilys

2.1 Introduction

This chapter examines some of the major steps involved in designing and undertaking a sampling program to assess exploited fish on coral reefs. Most of the discussion can also be applied to exploited invertebrates, such as trochus and beche-de-mer. There are two important things to consider in the early stages of a study: first, decide as precisely as possible what the major questions of interest are and plan how to address them; second, seek, wherever possible, the advice of other experts (also called peer review) to ensure that sampling is properly designed, implemented, analysed and interpreted. Good science relies on peer review to ensure the validity of the experimental design, appropriate interpretation of results and, ultimately, the best use of study resources.

This chapter provides guidelines for defining the questions of interest in regard to fisheries on coral reefs. It then provides a framework for addressing these questions by defining hypotheses that can be tested formally using statistical tests. Some of the background to these tests is then discussed, including the importance of considering the statistical power of tests. Different computer software programs for statistical analysis are also examined.

It is important to recognise that this chapter does not constitute a statistical text and it is assumed that fisheries biologists using the manual have a basic understanding of statistical testing or intend to seek further training in that area. Moreover, this chapter should be viewed as an introduction to some of the issues that must be considered in statistical analysis and provides some guidelines about sampling and how to apply some of the tests that are commonly used. For further reading, publications such as Green (1979), Snedecor and Cochran (1989), Underwood (1981, 1990, 1993, 1997), Andrew and Mapstone (1987), Fairweather (1991), Sokal and Rohlf (1981), Winer et al (1991) and Mapstone et al (1996) should be examined. Roger Green's (1979) book was a landmark in clarifying sampling design and statistical methods for environmental biologists, and remains highly relevant today. He summarised the correct approach to developing and executing environmental studies in his famous "Ten Principles", which are reproduced in Table 2.1. The AIMS Survey Manual for Tropical Marine Resources (English et al 1994) also provides good background information on the design and implementation of surveys in tropical marine habitats.

Table 2.1 TEN PRINCIPLES (source: Green 1979)	
1.	Be able to state concisely to someone else what question you are asking. Your results will be as coherent and as comprehensible as your initial conception of the problem.
2.	Take replicate samples within each combination of time, location, and any other controlled variable. Differences among can only be demonstrated by comparison to differences within.
3.	Take an equal number of randomly allocated replicate samples for each combination of controlled variables. Putting samples in "representative" or "typical" places is *not* random sampling.
4.	To test whether a condition has an effect, collect samples both where the condition is present and where the condition is absent but all else is the same. An effect can only be demonstrated by comparison with a control.
5.	Carry out some preliminary sampling to provide a basis for evaluation of sampling design and statistical analysis options. Those who skip this step because they do not have enough time usually end up losing time.
6.	Verify that your sampling device or method is sampling the population you think you are sampling, and with equal and adequate efficiency over the entire range of sampling conditions to be encountered. Variation in efficiency of sampling from area to area biases among-area comparisons.
7.	If the area to be sampled has a large scale environmental pattern, break the area up into relatively homogeneous subareas and allocate samples to each in proportion to the size of the subarea. If it is an estimate of total abundance over the entire area that is desired, make the allocation proportional to the number of organisms in the subarea.
8.	Verify that your sample unit size is appropriate to the size, densities, and spatial distributions of the organisms you are sampling. Then estimate the number of replicate samples required to obtain the precision you want.
9.	Test your data to determine whether the error variation is homogeneous, normally distributed, and independent of the mean. If it is not, as will be the case for most field data, then (*a*) appropriately transform the data, (*b*) use a distribution-free (nonparametric) procedure, (*c*) use an appropriate sequential sampling design, or (*d*) test against simulated H_o data.
10.	Having chosen the best statistical method to test your hypothesis, stick with the result. An unexpected or undesired result is *not* a valid reason for rejecting the method and hunting for a "better" one.

Sampling design

2.2 Defining the question(s)

Most, if not all, fisheries studies are based on a need to address a question or questions about observations made in nature (cf. Underwood 1990). These questions may arise in relation to an existing condition, for example we may ask: are current fishing practices having an effect on fish stocks on a certain reef? They may also arise in relation to a future condition, for example, if fishing is allowed - or if existing fishing methods change - on a certain reef, what will be the effect on fish stocks there?

It is important to recognise that in asking such questions there are implicit and pre-conceived theories which form the basis of the question. These may be derived from our knowledge of the effects of fishing on other reefs, or some intuitive logic (e.g. if fishing increases, finite stocks of fish should intuitively decrease). Alternatively the theories may be derived from a management agency acting in response to concerns by local villagers. The underlying basis, the theories, will often determine how the study is done and what components of the fish stocks and their environment are measured.

Since the initial question(s) play such a critical role in fisheries investigations, two steps are strongly recommended at the very start of a study:

(i) ensure all stakeholders (e.g. local communities, managers, collaborating scientists) have a clear understanding of the question(s) being addressed

(ii) ensure study methods and statistical procedures that will be used to answer the question(s) are identified

In far too many cases the questions are poorly defined and the methods and statistical procedures are inappropriate for answering the question(s) of interest. One way in which questions may be refined is by doing a small pilot study. The use of pilot studies is highly recommended, not only in helping to focus the aims of a study, but in refining sampling methods and determining the optimal sample size (see Chapter 3 and English et al 1994).

2.3 Creating a logical framework for answering scientific questions

The process by which scientists go about addressing scientific questions has received considerable attention and the methods developed are as applicable to fisheries science as to any other branch of science. The following discussion is drawn from two key references: Green (1979) and Underwood (1990).

Underwood (1990) identified the logical components in what has become known as a *falsificationist* or *refutationist test*, so-called because the emphasis is on disproving an hypothesis rather than proving it (see Underwood 1990 and references therein for more discussion of this). The general framework is shown in Figure 2.1 and examples developed in relation to coral reef fisheries are shown in Figures 2.2 and 2.3.
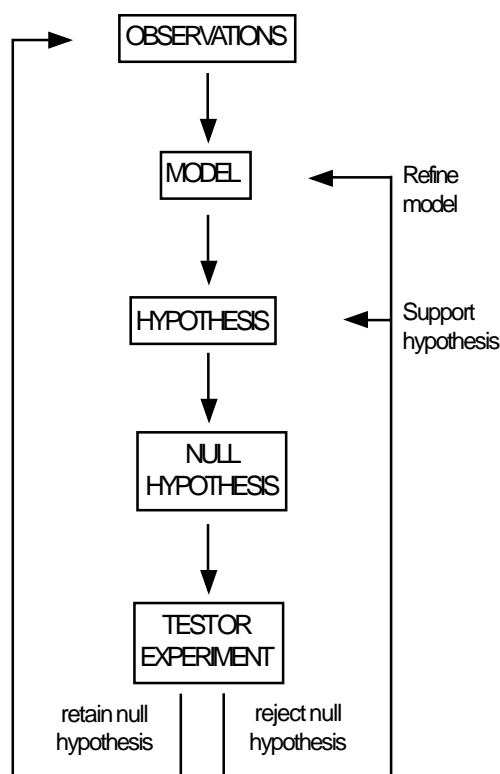
Figure 2.1 The logical components of a falsificationist experimental procedure. Source: Underwood (1990)

The process (Figure 2.1) starts with *observations* from nature.

The observations can also be considered as puzzles or problems (Underwood 1990) that may have been identified by others.

The *model* is simply a statement providing an account or explanation of the observations.

A model attempts to put forward theories to provide a realistic explanation of the observations, based on currently available information. As Underwood points out, however: "...articulation of a model is insufficient to demonstrate its

validity and some procedure is needed to contrast or compare different models that can be proposed to explain some observation." (Underwood 1990, p. 367).

An *hypothesis* is then proposed that can be tested (Figure 2.1). The hypothesis is a prediction about the model in relation to some new, as yet unexamined, set of observations (Underwood 1990). *It is crucial to recognise that one cannot use data to propose an hypothesis and then use those original data to test the hypothesis.*

Setting out to prove the hypothesis requires the use of a *null hypothesis*, which is the logical opposite statement to the hypothesis. It is used as a disproof device and includes all possibilities other than the prediction of interest. The next step in the process is the evaluation of the null hypothesis, which often takes the form of a *test* or *experiment*.

Underwood (1990, 1997) provides clear and detailed explanations of this procedure. The reason we use the null hypothesis is that it is impossible to prove a hypothesis because proof requires every possible observation to be available. i.e. we would have to assume that what happens from the cases observed in our test occurs in all possible circumstances. The use of the null hypothesis is known as the falsificationist procedure, because we attempt to disprove the null hypothesis rather than prove the hypothesis. Disproof of the null hypothesis, by definition, leaves the original hypothesis as the only alternative (Underwood 1997, p15).

Note that the use of a statistical test is not an inherent part of the process described above. It does, however, provide an objective means of evaluating the new observations (i.e. data) obtained to test the predictions of the model (see below). As shown in Figure 2.1, the outcome of the test will provide an indication as to the future direction of research, either in terms of evaluating the observations and developing another model (if the null hypothesis is retained); or refining the model to investigate if the model holds under different conditions, etc (if the null hypothesis is rejected). This latter approach is similar to the concept of adaptive management suggested by Walters (1986) in that the outcomes of the falsificationist test can be used to refine management practices.

Figures 2.2 and 2.3 provide an example of how we might
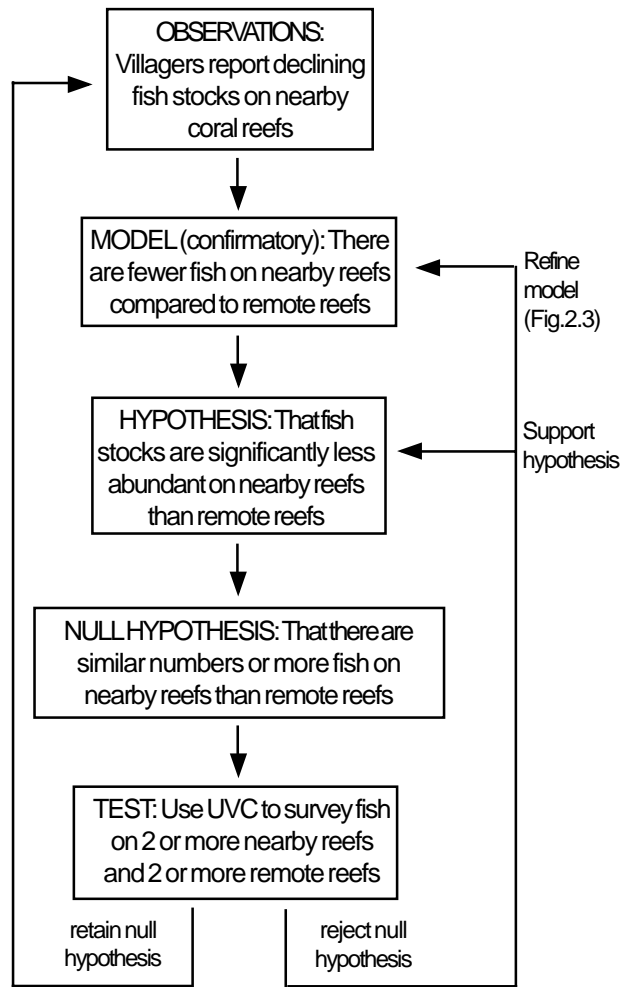
Figure 2.2 Use of a confirmatory model to investigate reported declines in fish stocks in the vicinity of an island village.

apply the above approach to coral reef fisheries. Initially, local villagers on an island report to their Division of Fisheries that catches of fish have declined on reefs close to their village but that catches remain large on reefs a long way from the village. Fisheries Officers are required to evaluate this claim. The first step is determine the likelihood that there are, indeed, lower catches from the nearby reefs than remote reefs. Here, the model is essentially confirmatory and the hypothesis predicts that collection of data from nearby and remote reefs will indicate that stocks are lower on the nearby reefs. Consequently, the null hypothesis is that stocks are either the same or greater on nearby reefs compared to remote reefs (Fig. 2.2). The test of the null hypothesis is a UVC survey of several reefs near to the village and several reefs in remote areas. Note that if we were to sample only one reef within each location, we would not know whether any differences detected were due to a general condition within the location (which *might* be due to fishing pressure) or to some unique ecological attribute of the reef (for example limited recruitment). Here

we would say that the two effects: 'general location' and 'uniqueness of the reef' were **confounded**. Another term used for this problem is **pseudoreplication** (Hurlbert 1984). The way we address this issue is by sampling at two or more reefs within the location, so we can obtain a measure of the average population size in that location. As the number of reefs within locations increases, so too does our confidence about making general conclusions regarding the location close to the village. Ideally, sampling should be done at four to six reefs to give us a confident measurement of the condition of fish stocks in the location.

In the example discussed in relation to Fig 2.2 our test may be a nested analysis of variance comparing locations (nearby vs remote) and sites within location (technical details of this test are discussed in Chapter 6). If the test for locations were significant, and the test for sites significant or non-significant, we would inspect the means to determine if the mean abundances were less near to the village. If so, we would reject the null hypothesis and confirm our model.

If the test for locations and sites was not significant we would retain the null hypothesis and conclude that the model was not supported by our observations. In this case we would either reject the villagers assertions and/or seek further observations from the villagers (e.g. regarding weather conditions at the time of fishing or possibly some social factors) that may lead to another model.

Finally, if the test for locations was non-significant but the test for sites was significant, this would suggest variability in fish stocks at a smaller spatial scale than locations. These observations may lead us to define another model based on the scale at which an effect of fishing might be apparent.

Assuming that we reject the null hypothesis, we can conclude that we have identified a pattern, i.e. that fish stocks are lower on nearby reefs compared to remote ones, but we have not unambiguously demonstrated that the cause of the difference is due to fishing. There are likely to be numerous possible alternative explanations, such as impacts related to runoff from agricultural practices, resort development, or possibly some natural factor. Our work would, however, allow us to refine our model which could explain the pattern observed. This new model is shown in Figure 2.3 and it leads to a hypothesis that predicts that if we reduce fishing pressure on some nearby reefs we will observe an increase in stocks there compared to nearby fished reefs. We might also hypothesise that abundance on nearby unfished reefs would approach and possibly exceed that of the remote reefs. Our null hypothesis would be that reduction in fishing pressure on nearby reefs would have no effect on stocks there compared to nearby reefs that are fished. Our test would be to do at least two UVC surveys on at least four reefs near the village, then close half of all the reefs to fishing for an appropriate period of time to allow stocks to increase. We would then do at least two more UVC surveys on all reefs. This type of test is being done currently in relation to exploited species of invertebrates on coral reefs in Solomon Islands (Lincoln Smith and Bell 1996).

If the null hypothesis is rejected, the study findings could be used for adaptive management to regulate fishing pressure on reefs close to the village. If the null hypothesis was accepted, we would look for another model to explain our observations (Fig. 2.3).
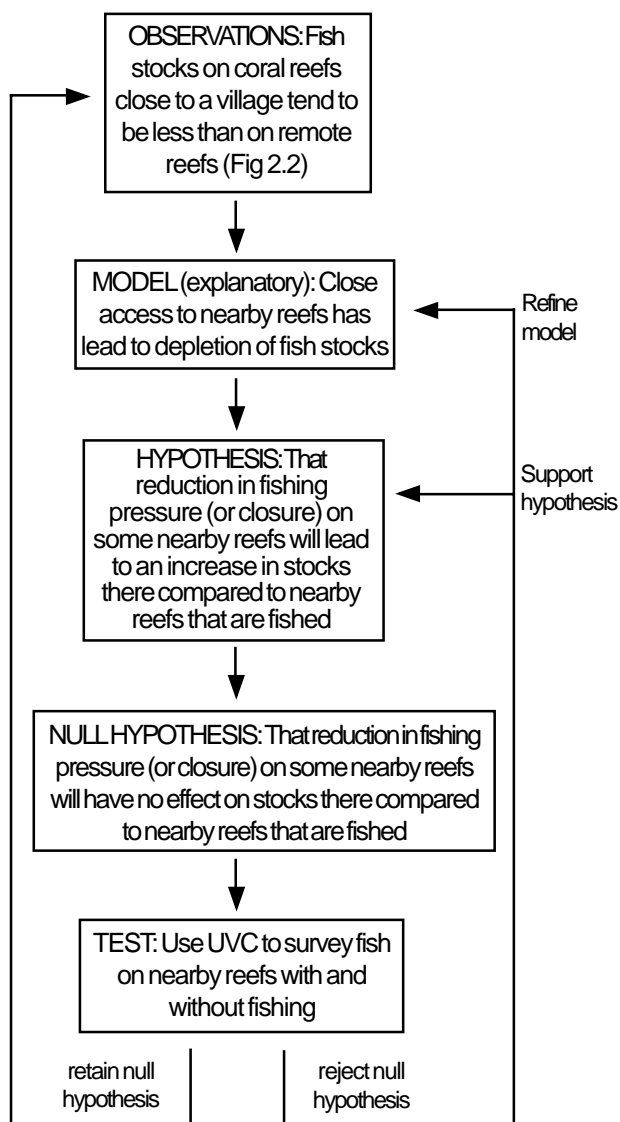
Figure 2.3 Use of an explanatory model to investigate reported declines in fish stocks in the vicinity of an island village

2.4 An introduction to statistical tests

2.4.1 What is biological sampling?

Nearly all collection of data in ecological studies requires sampling because we cannot directly count the total population of any species we may be interested in (see Zar 1984 p.16). For example, if we wish to know the population of coral trout on a large reef it would be very difficult to count all individuals. In fact, the tests described in the previous section would all rely on sampling to provide an indication of fish stocks on fished and unfished reefs.

Sampling means that we take standardised, representative measures of the species of interest from the site(s) of interest. For example, UVC provides counts of fish within clearly defined areas of reef that we can count manageably (Chapter 3). By taking a number of units (counts) - usually called *replicates* - in different patches of reef we can obtain one sample made up of several *independent* measures of the density of fish. These terms are further explored in Chapter 3. Note that taking several counts of fish in exactly the same patches of reef does not provide independent replicates, which are a crucial prerequisite for the statistical testing recommended in this manual. In designing a sampling program we consider how best to sample the population to obtain as *accurate* and *precise* an estimate of the total population as possible. Accuracy and precision relate closely to the sampling methodology (Chapters 3 and 4), sampling design, and the statistical tests employed to test the data. *Accuracy* refers to how close the estimate comes to the true value. *Precision* refers to the spread or variation in the data. Andrew and Mapstone (1987) provide an excellent review of these terms and their importance in designing sampling programs.

By obtaining a number of replicates from a site of interest we can obtain an estimate of the *average* or *mean* density of fish and of the *variance* associated with that mean. The variance is a measure of the dispersion or spread of the data and it can be used to calculate a number of statistics, including *standard deviation* and *confidence limits*. These terms are important in statistical testing and are discussed in more detail in Chapter 6.

Once we obtain a mean estimate of, for example, coral trout density, we can determine the population of the entire reef by multiplying the total area of the reef (which may be estimated from admiralty maps, aerial photos, etc) by the mean density to obtain an estimate of the total abundance.

Similarly, we can multiply our confidence limits by total area to estimate the likely range in total abundance. In many cases, however, we are not particularly interested in total abundance on a reef, but use the relative abundance (i.e. the size of the mean and its variance) to compare among this and other reefs, or to compare the same reef at different times.

2.4.2 Why do we use statistical tests?

Statistical tests are an essential part of ecology and their use in the last two decades has become extremely widespread and often highly sophisticated. Statistical tests are also becoming far more common in surveys of coral reef fisheries: they are a means of objectively evaluating information collected about the impacts of humans on the environment and fisheries. Statistical tests are based on the notion of determining the likelihood, or probability, that data collected are consistent with a pre-determined hypothesis or question (e.g. that populations of a species are less abundant, on average, at one site than at others).

By convention, scientists give themselves a 5% chance of accepting that there was an hypothesised effect when in fact there really wasn't one (see 2.4.4). Apart from being relatively objective, the great strength of statistical design is that, if done properly, it compels researchers to collect their data within a logical framework to address specific questions of concern. Moreover, the more precise the question, the more likely we are to obtain an unambiguous result (i.e. there was a difference or there wasn't). One potential problem of statistical testing is that it is often difficult to present findings concisely to local communities and managers. It requires considerable effort to ensure that statistical findings are made comprehensible to all stakeholders.

Notwithstanding their potential complexity, a statistical test allows researchers to assess if differences observed from sampling are likely to represent true differences between conditions or situations (e.g. times, sites, fished vs unfished, etc - also generally called *factors*, *effects* or *treatments*) being compared, or merely reflect a chance effect (Manly 1991). A critical step in the process is the definition of hypotheses that are to be tested. Green (1979) and Underwood (1990) provide a good background to the logics of statistical testing in ecology and this can be readily extended to fisheries investigations, including UVC and CPUE surveys.

2.4.3 Selection of tests

Many ecological studies use two basic kinds of statistics to evaluate the aquatic environment: *univariate* and *multivariate* statistics. Within each of these, there are *parametric* and *non-parametric* tests. Parametric tests are based on measures of central tendency (usually the mean) and dispersion (usually the standard deviation) and make assumptions regarding the distribution of the data (usually assuming a normal distribution). Non-parametric tests are often based on ranks or proportions which do not assume an underlying normal distribution of the data. By-and-large, parametric tests are more powerful, can be used to evaluate highly complex or multifactorial questions (see below) and, thus, tend to be preferred. Recently, statisticians have developed computer-intensive *randomisation* or *permutation* tests which compare a test statistic for the sample data against a distribution created by randomising the sample data many times and re-calculating the test statistic each time (Siegel and Castellan 1988, Manly 1991). Whilst rarely seen in past studies of tropical reef fisheries, these tests are becoming increasingly more common. The following section provides a general introduction to statistical tests. The application of these tests is described in further detail in Chapter 6.

Univariate tests

Univariate tests examine hypotheses related to a single dependent variable in relation to one or more independent variables. A departure from this is correlation analysis, where variables compared may be dependent on each other or they may be dependent upon some other variable. Dependent variables can include counts of fish, fish sizes, weights, etc. In addition, dependent variables often include "derived variables", which are measures synthesised from the sample data. Examples of dependent (or "derived") variables include total abundance (i.e. individuals of all species within a sample), species richness (i.e. the number of species within a sample) and community indices (e.g. diversity, evenness and similarity measures). Independent variables can include factors such as location, time and tide state; or they may represent some experimentally-varied factors such as type of gear (e.g. trap size, hook size), which, under experimental conditions, are varied by the investigator. In tropical fisheries, human activities (e.g. line fishing, spearfishing, coastal development, etc) may be seen as experimental conditions potentially affecting a number of dependent variables (Carpenter

1989, Lincoln Smith 1991, Underwood 1995).

Parametric univariate tests commonly seen in UVC studies include t-tests, correlation, regression and analysis of variance (ANOVA). Another class of tests commonly used are goodness-of-fit tests, including the Chi-squared test, used to compare the observed proportions of a dependent variable against what might be expected by chance alone. Descriptions of these tests are provided in a number of texts (e.g. Snedecor and Cochran 1989, Sokal and Rohlf 1981, Siegel and Castellan 1988 and Winer et al 1991). Underwood (1981, 1997) provides a detailed synthesis of the use of ANOVA in marine ecology.

The selection of univariate tests to examine hypotheses requires careful consideration. Moreover, the use of parametric tests requires that the assumptions underlying their use are tested. Violation of some of the underlying assumptions can be mitigated by transforming the data (e.g. to a logarithmic scale) or by conservative interpretation of the results (e.g. by reducing the acceptance level from 5% to 1%; or, for some questions, by increasing it to say, 10% - see below). Notwithstanding this, failure in properly designing programs for data collection, or using tests inappropriately, can lead to false conclusions with potentially costly consequences. Underwood (1981) provides a good discussion of the assumptions that must be met for ANOVA.

Multivariate tests

Multivariate statistics include a large variety of procedures which essentially cluster groups of variables according to their similarity or dissimilarity (Field et al 1982, Faith et al 1991, 1995, Clarke 1993). When originally developed, they were used for inferring patterns or *generating hypotheses* without a rigorous framework for *hypothesis-testing* (see above). More recently, both parametric and non-parametric procedures for testing hypotheses in multivariate statistics have been developed. Parametric tests, including multivariate analysis of variance (MANOVA) are often avoided because of difficulties with satisfying the underlying assumptions of the test (Johnson and Field 1993). Non-parametric procedures called ANOSIM (analysis of similarities) have, however, been developed based on randomisation tests (Field et al 1982, Clarke 1993). While ANOSIM procedures are applicable to a wide variety of data sets, they are currently limited to more simple designs than are being evaluated using univariate tests such

as ANOVA. Notwithstanding this limitation, multivariate analyses, including ANOSIM, are valuable because they allow us to test hypotheses about variation at the level of assemblages of fish.

In aquatic ecology, multivariate analyses are applied to samples containing an assemblage of fish or invertebrates, often analysed at the species or family level and used to compare locations and/or times of interest. They have also been used with habitat variables to identify how the habitat characteristics of sites may explain differences in populations of exploited animals (Lincoln Smith and Bell 1996). A recent extension of multivariate analyses has been the development of SIMPER analysis (Clarke 1993), which indicates those taxa within an assemblage which contribute most to the dissimilarities between the factors of interest (e.g. sites). Such analyses can also be used to compare data across different sampling procedures. For example, Samoilys et al (1995) used multivariate analyses to compare the relative abundance of fish reported in creel and questionnaires surveys to the relative abundance of fish as estimated by UVC, on reefs in Fiji and Solomon Islands.

To summarise: although the question or hypothesis of interest will determine the type of statistical procedure used, fisheries scientists should consider using both univariate and multivariate statistics to examine data sets collected as part of a fisheries stock assessment. This approach allows an assessment of variability for *fish assemblages* (also referred to as *fish community*. i.e. how does the group of species sampled varied as a whole?) and for populations of species within the assemblage. The former may become increasingly important in multispecies fisheries assessment as the preferential removal of some groups of species (e.g. piscivores such as Serranidae, Lutjanidae and Lethrinidae) probably causes changes in the structure of assemblages (Jennings and Lock 1996). The latter is particularly important if we are concerned about the response of a particular species to fishing. When examining the fisheries resources (and habitat characteristics) of sites of interest, it is often useful to use both univariate and multivariate statistical procedures to evaluate variation at the level of populations and assemblages, respectively.

2.4.4 The power of statistical tests

In using statistical testing in fisheries science, it is possible to be correct in two ways or incorrect in two ways. One may be correct in concluding that an effect (or "difference")

occurred and in reality it did; or it is possible to be correct in inferring that no effect occurred when there was no impact (e.g. from fishing). Alternatively, we may incorrectly conclude that an effect was present when in fact there was no effect. This would happen when the probability of the test statistic was equal to or less than 0.05 (i.e. $P \leq 0.05$, or whatever acceptance criterion we selected prior to doing the test), but the sample data did not truly reflect the condition in nature. Being wrong in this way is generally denoted as a *Type I error* and the probability of making this type of error is symbolised by alpha (α). On the other hand, we may incorrectly conclude from our study that there was no effect, when in fact there was. This would happen when $P > 0.05$, or some other acceptance criterion. This type of mistake is generally called a *Type II error* and the probability of making this type of error is symbolised by beta (β). Table 2.2 summarises these four possibilities.

TABLE 2.2 The two types of errors in hypothesis testing (Source: Zar 1984)

	If H_o is true	If H_o is false
If H_o is rejected:	Type I error	No error
If H_o is not rejected:	No error	Type II error

Arising from these alternatives is the notion of statistical power, which basically asks: how effective is the sampling program at answering the question of interest? Or, more formally, the power of a statistical test (1-β) is the probability of rejecting the null hypothesis when it is false and thus, should be rejected (Siegel and Castellan 1988).

The concept of statistical power is fundamental to the use of statistical testing in surveys of fisheries resources using UVC, creel surveys, etc. In considering how to use power analysis for fisheries research, the following three points are noteworthy. First, the concept of statistical power can be considered in terms of risk to the environment. Thus, one may argue that it is better, from the point-of-view of maintaining the fish stocks, to commit a Type I error (i.e. to conclude that there was an effect or a difference when in fact there was none) than a Type II error (i.e. to conclude there was no effect or difference when in fact there was - Table 2.2). If this view is adopted, we may increase the acceptance criterion from 0.05 to say, 0.10 to reduce the chance of a Type II error (see below) and this may be an

appropriate approach if the cost of an impact is very high (e.g. loss of an important fishing ground).

It is important to recognise, however, that the particular approach adopted can lead to increased and possibly unnecessary hardships to those whose fishing practices may be limited (e.g. if an effect is incorrectly inferred) or, alternatively, to the environment and possibly future generations (e.g. if no effect is incorrectly inferred). These issues have been discussed by Underwood (1993) and particularly by Mapstone (1995).

Second, researchers have some scope for varying the power of a statistical test. Power is affected by the sample size used, thus collecting more samples increases statistical power. It is also affected by the acceptance criterion as discussed in the previous paragraph, but this has the drawback of increasing the potential for committing a Type I error. Power is also affected by the extent of variability in the system being studied, thus large variability leads to low power. This factor cannot be controlled by the researcher other than by trying to maximise sample sizes and possibly by rejecting from surveys some species that require huge sample sizes to be able to detect differences. Finally, statistical power is affected by the size of the difference (or effect) that may be considered important. For example, we might specify that a 40% decrease in the catch of coral trout is something to be concerned about. As the "effect size" increases, so does statistical power. Determining effect size should be an important part of the scoping phase or pilot study of a fisheries investigation.

Third, power analysis can be used in two broad ways. It can be used to evaluate a study program that has been completed (i.e. how confident can we be in the conclusions drawn from statistical testing, particularly where non-significant results were reported?). Alternatively, it can be used to design further studies, by helping with selection of sample sizes, effect sizes and decision variables that are cost-effective. Fairweather (1991) provides a good discussion of the uses of power analysis in aquatic ecology, other references include Underwood (1981), Cohen (1988), Peterman (1990), Mapstone (1995) and Mapstone et al (1996).

2.5 Why is replication so important and what is the optimal number of replicates that should be collected?

As implied from the foregoing discussion, the collection of replicates is a major consideration in the design of sampling programs. Obtaining a sample of replicate units enables the calculation of means and variances which form the basis for estimating the size of stocks and for most parametric tests. Fundamentally, replication prevents us from confounding variability associated with a single measurement with the treatment (e.g. reef) we are interested in comparing. This notion should be considered for all treatments that are examined as part of a survey. Hurlbert (1984) and Stewart-Oaten et al (1986) provide detailed discussions of the consequences of failing to replicate at all levels of interest.

Having emphasised the need for replication in sampling fish stocks, the next task is to determine how many replicates (i.e. the sample size) are required to give us a good chance of detecting the hypothesised effect. For UVC on coral reefs, extensive work has already been done on the amount of replication required and this is a good basis for the design of future studies (see Chapter 3). However, even though there is a good basis for determining replication for UVC, it is still important to evaluate if this is sufficient for particular studies. Discussions of how to select the optimal sample size is provided by Green (1979), Sokal and Rohlf (1981), Andrew and Mapstone (1987) and Bros and Cowell (1987).

2.6 Statistical software

There is a large variety of computer programs available to do the types of statistical analyses required for fisheries studies. Some of the programs commonly used include SAS, SPSS, MINITAB, Systat and Statistica for univariate and multivariate parametric and non-parametric tests; GMAV5 for analysis of variance; and PATN and PRIMER for multivariate statistics. Some of the spreadsheet and database programs can also be used for statistical testing, particularly randomisation tests, although they are sometimes limited in the number and complexity of tests available for parametric tests such as ANOVA. Computer programs required for data storage and manipulation are discussed in Chapter 5.

When using statistical computer programs it is extremely important to know how the data are being treated by the program and to order the data appropriately so that the program reads columns and rows correctly. For example, in analysis of variance it is important to specify whether factors are fixed or random; or nested or orthogonal (see examples below). Failure to do so will lead to default settings being used which may provide an incorrect result for the design being used. Also, some programs will analyse unbalanced or un-replicated data sets. If such data sets must be used, it is essential that the underlying assumptions and models used by the program are understood.

When using an unfamiliar computer program for statistical analysis it is highly desirable to repeat analyses that have been done on more familiar programs using the new program to check that the same result is obtained. Alternatively, many statistics texts (e.g. Winer et al 1991) provide worked examples of tests with raw data which can be used to evaluate a program. Finally, researchers should graph their data (usually summarised as means and standard errors) to see if the outcome of the test is consistent with the graphical interpretation of plots (see Chapter 6).

In summary, there is a wide variety of computer programs available for handling most statistical analyses required in fisheries science and statistical hypothesis testing.

When using a new program, carefully evaluate data input and test outputs of the software. Check new programs by running data sets with known outcomes and compare the test results with plots of the data to ensure consistency.

CHAPTER THREE: UNDERWATER VISUAL CENSUS SURVEYS

Melita Samoilys

3.1 Introduction

Underwater visual census (UVC) is a technique commonly used to measure the abundance of fishes on coral reefs, and has been used extensively in reef fish studies of population dynamics, ecology and management (see reviews by Barans and Bortone 1983, Harmelin-Vivien et al 1985, Thresher and Gunn 1986, Cappo and Brown 1996). UVC has also been used to census a wide range of species that are taken by shallow water demersal fisheries on coral reefs (Russ 1985, Kulbicki 1988, Samoilys 1988, McManus et al 1992, Ayling & Ayling 1992, Roberts & Polunin 1993, Watson & Ormond 1994, Samoilys et al 1995, Jennings & Polunin 1996). Visual census methods can provide rapid estimates of relative abundance, biomass and length frequency distributions of reef fish. UVC methods allow researchers to focus on key species of particular relevance, are non-destructive and, unlike most fisheries data, collect fishery-independent data on stock abundance. UVC methods are usually done using SCUBA, though sometimes snorkel can be used in shallow habitats. Thus, they take the fisheries scientist into the water which encourages awareness of the environment and fish ecology, and provides an opportunity for detecting habitat impacts such as coral damage from siltation, dynamite fishing, etc.

The main disadvantage of UVC methods is depth constraints imposed by SCUBA diving, thus the full range of a species' distribution may not be surveyed. This issue should be considered when formulating questions and designing a research program. Other disadvantages include the restriction to species that are diurnal, visually obvious, and not repulsed by divers, and the potential for observer error and bias in estimating numbers and sizes of fish. The interaction between fish and divers has been demonstrated (Watson et al 1995) and remains a potential source of error in the visual estimation of population abundance.

A variety of UVC methods have been used (e.g. Thresher and Gunn 1986, and reviewed recently by Cappo and Brown 1996) ranging from strip transects, a method originally put forward by Brock (1954), to stationary point counts (Bohnsack and Bannerot 1986). This manual describes the procedures for doing both strip transects and stationary point counts based on the methods developed during the ACIAR/DPI UVC project (Samoilys and Carlos 1992, Samoilys et al 1995).

This chapter provides a general procedure for conducting UVC surveys applicable to most shallow water coral reef environments in the tropical Pacific. It is important to note the principles of the procedure so that if an unusual sampling situation arises an appropriate specialised sampling program can be designed based on the same principles. The AIMS Survey Manual for Tropical Marine Resources (English et al 1994, pp. 68-78) describes procedures for censusing a wide range of reef fish species using 50m x 5m strip transects. The procedures described in the present chapter are similar to those of the AIMS manual except here we focus only on food fish - species exploited in Pacific fisheries, and we also describe the stationary point count technique.

3.2 Design of surveys

The design of a UVC survey will depend on what questions are being asked about the population densities of reef fishes. For example do we want to compare populations between regions, between reefs, or between habitats? What species are we interested in, and what other factors are involved, such as fishing pressure, season, weather, impacts from agriculture, etc? Procedures and principles for defining questions and for designing surveys are detailed in Chapter 2.

> **1. Decide on the objectives of the survey, the scale of sampling and the design in terms of *strata* and levels of *replication*. Formulate the questions and indicate the tests that may be used. Strata may refer to factors such as fishing pressure, habitat type, spawning season etc. *Replicate sampling units* are located within strata, either spatially or temporally depending on the nature of the strata.**

A survey invariably involves different strata. Selecting strata depends on the questions being asked (see Chapter 2). These may refer to *factors* (also called *treatments*) such as habitat
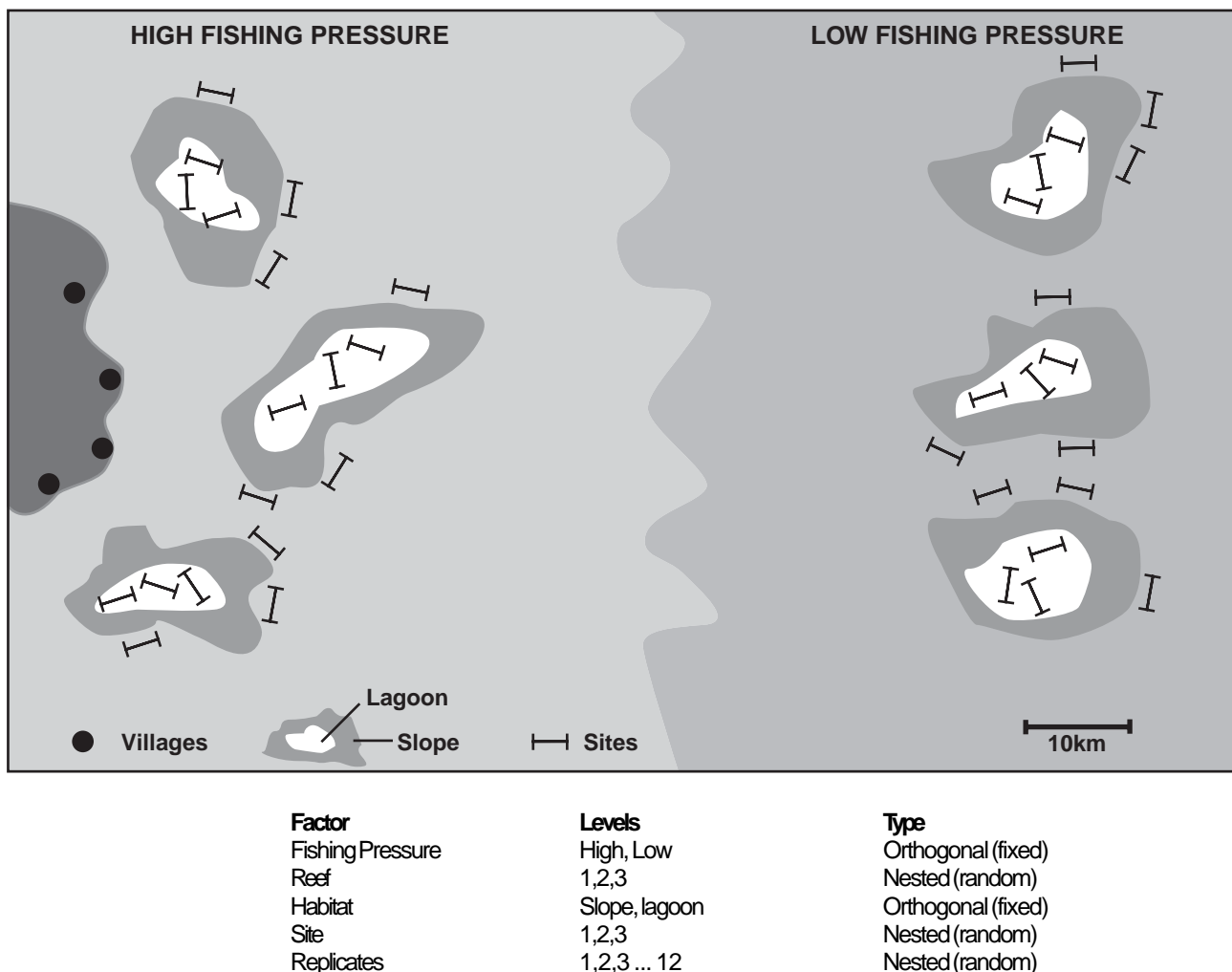
type, fishing pressure, distance from shore, spawning season etc. ***Replicate sampling units*** or ***replicates*** are placed within strata. Factors may be ***fixed (=orthogonal)*** or ***random (=nested)***. A design which has both fixed and random factors is called a mixed model in statistics. These terms are explained below. They relate to the questions being asked and the subsequent statistical analyses (e.g. ANOVA) that will be used (Chapter 6).

Figure 3.1 provides an example of a UVC survey design which is stratified according to fishing pressure and habitat type. These two factors have been identified by the researcher's questions. For example, we may ask: are fish more abundant on reefs that are lightly fished, and do fish densities differ between slope and lagoon habitats? In this case fishing pressure and habitat are fixed factors - they have been specifically selected for study and their characteristics identified. In Figure 3.1 two other factors are included: reefs and sites. Here, we have decided to sample the strata within discreet units - reefs - which are nested within fishing pressure. e.g. 3 reefs within a lightly

fished area and 3 reefs within a heavily fished area. The reefs have been selected randomly. They are random factors because our question relates to fishing pressure, not reefs. Any three reefs could be chosen. In addition, we suspect that populations of fish are likely to vary *within* each of the habitats within a reef. Therefore, we restrict the replicates to smaller areas or ***sites*** which are allocated *randomly* within each habitat. Sites are therefore a random factor, nested within habitat. The nesting or hierarchical aspect of this design enables us to look at what scale the variability in fish abundance occurs. It is also often logistically easier to sample within smaller areas. From a statistical perspective (e.g. using ANOVA, see Chapter 6), restricting replicates to sites is a more powerful way of looking for differences between the fixed factors - habitat and fishing pressure.

A hierarchical design can also be applied to sampling through time. For example, we may wish to examine how populations vary at different time-scales (e.g. between years, months within years or weeks within years and months). As with spatial sampling, there are important logistical

UVCSurveys

Factor	Levels	Type
Fishing Pressure	High, Low	Orthogonal (fixed)
Reef	1,2,3	Nested (random)
Habitat	Slope, lagoon	Orthogonal (fixed)
Site	1,2,3	Nested (random)
Replicates	1,2,3 ... 12	Nested (random)

Figure 3.1 A mixed model sampling design with orthogonal and nested factors, stratified according to fishing pressure and habitat type.

reasons why we may wish to apply a nested design to temporal sampling: it is often much easier to sample in manageable blocks of time within a year than to return to a site at randomly allocated times over a year.

The foregoing discussion illustrates an important philosophical difference between two objectives in surveying fish populations:
a) estimating total population size
b) detecting differences in population size

If replicates are restricted to sites (as illustrated in Figure 3.1), we are unable to obtain an unbiased estimate of total population size. In other words, if the total abundance of fish on a reef is required, replicates should be placed randomly *throughout* the habitats (strata) of the reef (e.g. McCormick and Choat 1987). However, if we are interested in comparing reefs, restricting replicates to random sites nested within reefs is logistically easier and also more powerful statistically. This chapter focusses on the hierarchical approach for these reasons and because objective (b) above is often more frequently required.

Stratified sampling is termed simple and random if equal numbers of replicates are allocated to each stratum. A more efficient design is optimal (or Neyman) stratified sampling which allocates different numbers of replicates to different strata because different strata may require fewer or more replicates depending on the variability in the data and the contribution of each stratum to the whole area. Optimal sampling is the most appropriate design for estimating total population size (objective (a) above). McCormick and Choat's (1987) study provides an excellent example of optimal stratified sampling. Hierarchical (nested) survey designs are more appropriate for detecting differences in population size (objective (b) above); they are discussed further in the application of statistical tests in Chapter 6.

2. A pilot study is strongly recommended: visit the study area and, with the help of aerial photographs, nautical charts and local information, define the strata and the sites. A preliminary assessment of the general area is required to select locations for the sites. This can be done on snorkel, or using manta boards. It is particularly useful to combine the UVC pilot study with the frame survey - the pilot study recommended for fishery surveys (see Chapter 4).

Sites within strata should be identical in dimension.

They should also be similar in physical characteristics, coral cover etc., i.e. they should cover a *homogeneous* area of habitat and not cross habitat boundaries. If different habitats are of interest then habitat is specified as one of the strata and replicate sites are located within each habitat. Sites should be separated by at least 100 - 200m. The exact positions of the sites should be recorded, either by taking bearings or, if available, by GPS (global positioning system).

Sites are selected *randomly* as representative areas of the general location being studied. For example two or three sites may be established along one side of a reef. We establish *replicate* sites because in choosing only one site we may have inadvertently selected a rather unrepresentative area, and therefore *biased* the results. By having more than one site we help avoid bias. Also, three or more sites greatly improves the statistical power of the design (see Chapter 6), which in turn enables us to make more general conclusions from the results about the reef. Put another way: by restricting replicate sampling units to discrete sites, variability is *partitioned*, which is more powerful statistically because the variance in the data associated with site differences can be identified. For example, it is more powerful statistically to sample the side of a reef with 10 replicate sampling units in two sites, than 20 replicate sampling units along the reef. This example shows that in terms of effort the use of two sites does not necessarily increase our field effort. These principles are further discussed in Chapters 2 and 6.

The reef may be rather patchy. If this is the case it is important to think about minimising the bias in choosing the sites - they should be representative of the general area, but not necessarily the best areas for finding lots of fish. If the habitat is very patchy, for example in lagoonal areas with large areas of sand, then sites should be located where there is coral since the UVC surveys are focussed on reef-associated fishes. It is important to note such details when writing the methods and analysing the results - the fish counts will relate to areas of coralline habitat rather than sandy habitats.

QDPI/ACIAR UNDERWATER VISUAL CENSUS DATA SHEET

Observer One:

Date: Reef: Site: Transect:

Time: Tide: Sea: Vis: Depth:

SCARIDAE	I. Phase	T. Phase	LETHRINIDAE	
Bolbometopon muricatum			Gymnocranius sp.	
Cetoscarus bicolor			Lethrinus harak	
Hipposcarus longiceps			L. nebulosus	
Scarus altipinnis			L. olivaceus	
S. rubroviolaceus			L. xanthochilus	
S. ghobban			L. spp.	
S. microrhinus			Monotaxis grandoculis	
S. niger				
S. frenatus				
S. spp.				
LUTJANIDAE			**SERRANIDAE**	
Aprion viriscens			Anyperodon leucogrammicus	
Lutjanus argentimaculatus			Cephalopholis argus	
L. bohar			C. cyanostigma	
L. carponotatus			C. miniata	
L. fulviflamma + L. ehrenbergi			Epinephelus caeruleopunctatus	
L. fulvus			E. maculatus	
L. gibbus			E. polyphekadion	
L. kasmira + L. quinquelineatus			Plectropomus aerolatus	
L. monostigma			P. laevis	
L. rivulatus			P. leopardus	
L. russelli			P. maculatus	
L. semicinctus			P. oligacanthus	
Macolor spp.			Variola spp.	
ACANTHURIDAE			**LABRIDAE**	
Acanthurus D-M-X			Cheilinus fasciatus	
A. lineatus			Ch. trilobatus	
A. nigricauda			Ch. undulatus	
A. triostegus			Choerodon spp.	
Ctenochaetus striatus			**NOTES**	
Acant. - Ctenot. - Zebras				
Naso brevirostris				
N. hexacanthus				
N. tuberosus				
N. unicornis				
N. spp.				

UVCSurveys

3. Define the species of fish to be studied, ensuring all species are suitable for UVC. The number of species should be minimised - the fewer selected the more accurate the counts will be. This step involves a compromise between the information required and the accuracy of the population estimates.

For further discussion on the issue of counting several species simultaneously see Lincoln Smith (1989). There are important criteria to consider when selecting species of fish for underwater visual census surveys. The fish should be:

a) highly visual and not cryptic

b) diurnal

c) not significantly underestimated by UVC (see Samoilys and Carlos 1992).

d) identifiable to species level (unless not required - e.g. only to genus, but this is not recommended).

A list of 60 species that are suitable for UVC surveys, as assessed by the ACIAR/DPI UVC Project (see Samoilys and Carlos 1992) are given in the sample datasheet. In general the list reflects those species which satisfy the criteria listed above, and are species that contribute to Pacific coral reef fisheries. Samoilys and Carlos (1992) should be consulted for species-specific details. The determination of accuracy of UVC for some species (e.g. some lethrinids, lutjanids and serranids) was inconclusive due to limited data probably caused by patchy distributions, low densities and diver-fish interactions (Samoilys and Carlos 1992). Nevertheless it is recommended that such species be included in the species list (if of interest to the researcher) since they would not substantially increase the cost of surveys and may provide broadscale density or presence/absence data (Mapstone and Ayling 1993). Note that some species are grouped in the species list because they are not easy to differentiate underwater. Thus *Acanthurus* D-M-X comprises *A. dussumieri*, *A. mata* and *A. xanthopterus*.

4. Define the size of sampling units. *Sampling units are the individual visual censuses. They represent the smallest sampling unit that is used to collect the data. For example a 7m radius point count, and a 50m x 5m strip transect are sampling units - these are some of the most commonly used dimensions.*

The size of the sampling unit is one of the first criteria to be considered when designing a UVC program. The size of a visual census count relates to the size of the animal being sampled and its range of movement. A general rule of thumb is that the ratio of the area (or volume) of the organism to the area (or volume) of the sampling unit should be negligibly small: 0.05 or less (Green 1979). In the case of mobile animals, such as fish, where observer avoidance may be a problem, the area of avoidance should be considered rather than the area of the fish (Green 1979). For example territorial pomacentrids will require smaller sampling units than the larger mobile lethrinids. Similarly, surveys of juveniles should use smaller units than surveys of adults (English et al 1994 (AIMS manual) p.86). Studies that have evaluated transect dimensions concluded that 50m x 5m transects were the most suitable for the larger species (fish >11cm FL) typically exploited in coral reef fisheries (Samoilys and Carlos 1992, Mapstone and Ayling 1993). This dimension is used by the AIMS coral reef monitoring team (English et al 1994: AIMS manual pp.68-78). Similar studies on stationary point counts concluded that a 7m (Samoilys and Carlos 1992) to 7.5m (Bohnsack and Bannerot 1986) radius count is the most suitable.

5. Define the number of sampling units. Based on previous work 10 replicates per site should be used for 50m x 5m transects and 12-16 replicates per site for point counts. If a pilot study (see below) cannot be done a sample size of at least 10 is strongly recommended.

These replication levels were determined for many species of reef fish that are important in the artisanal and subsistence fisheries of Fiji in Phase 1 of the ACIAR/QDPI Project (Samoilys and Carlos 1992). Twelve replicates were used for subsequent surveys using point counts in Fiji, Solomon Islands and Papua New Guinea (Samoilys et al 1995). A minimum sample size of 10 is recommended based on statistical considerations such as degrees of freedom and resolving power, which are particularly relevant to the highly variable distribution and densities typical of reef fishes.

Replication level can be evaluated by a quick and simple pilot study. This is recommended if different species and/ or habitats are being considered. Consider those factors that will determine how many replicates can be done and then select the maximum number of replicates that are feasible (e.g. 20 or 30 replicates). Collect count data using

this maximum number of replicates in a representative area to be studied. Plot the mean standard error (SE) against the sample size (n). The relationship is a decreasing asymptotic function approaching zero (Figure 3.2).

The number of replicates or sample size is an essential component of any experimental design (see Chapter 2). If the sample size is too small the power to detect differences between means is likely to be very low or inadequate, and if the sample size is too large effort is wasted. Bros and Cowell (1987) discuss these issues and describe a method for determining optimal sample size by defining the maximum sample size and the minimum sample size. The maximum number of replicates is based on factors such as time, money, materials and feasibility (Green 1979). The minimum number of replicates is defined in terms of resolving power (i.e. the power to detect change in fish abundance). The minimum acceptable sample size should be beyond the region of maximum change in the slope of the variability of the density estimates (see Figure 3.2). In other words there will be no appreciable improvement in power if sample size is greater than at this point on the graph, therefore the extra effort is not worth it. With patchy (clumped) distributions as is invariably the case with fish counts, the point at which the rate of change in the coefficient of variation of the density estimates is sharply reduced, may be used as the minimum acceptable sample size (after Bros and Cowell 1987). Plotting variability

functions requires bootstrapping (Samoilys and Carlos in prep).

The number of replicates or sample size at which the rate of change in the coefficient of variation of the density estimates is sharply reduced (as the curve begins to asymptote) is the replication number, n, that should be used. If there is no asymptote the results suggest the species, sampling units and/or study area selected will not give good data on fish populations. If this is the case, the design stage should be re-evaluated.

6. Decide on the duration of a census. For example the 7m radius point count used for counting ~ 60 species (see sample datasheet) was standardised to 7 minutes. If only a few species are being counted (e.g. 10) then only 2-3 minutes may be necessary.

A pilot study is useful: trial the selected species list and plot the cumulative number of fish against time (see Figure 3.3). The graph will asymptote. This represents the time at which all fish have been counted. Trials should be conducted in areas where fish are most abundant and/or the habitat is most complex, because these will require a longer count duration.

The duration of each census needs to be standardised because there are biases associated with the time spent in

UVCSurveys

Coefficient of Variation

Replication

Coefficient of Variation

Replication

Figure 3.2 Change in variability of estimates of mean density over a range of replication levels; data derived using coefficients of variation from bootstrapping density estimates from 30 point counts (10m radius) and 16 transect counts (50x5m) conducted in the same area of reef (Samoilys and Carlos In prep.).

the census area (e.g. fish attracted to or repelled by the divers, see Samoilys and Carlos 1992, Watson et al 1995). The time should be the minimum required to search the census area completely, since the longer the census the greater the problems of interference from divers, incoming fish (see below), etc. This issue is clearly discussed by Lincoln Smith (1988). When conducting a visual census the researcher is attempting to simulate an ***instantaneous "snap-shot" count***. i.e. in zero time. In reality this is not possible because it takes a finite amount of time to search the census area and count the fish.

3.3 Training in fish size estimation

Visual census counts for stock assessment purposes involve the visual estimation of fish sizes. Fish lengths are estimated to provide a size frequency distribution for the population, and to obtain biomass or weights of fish using length-weight relationships. Biomass is usually a more useful parameter in fisheries stock assessment. For example, yields are usually expressed in kg/km^2 (see Chapter 6). Fish length estimation requires training, and when counts are conducted over long periods of time, observers should also re-train or practice since they will lose the ability to estimate fish lengths accurately. Observers are trained to determine how accurate they are and to ensure that they are consistent.

Training is conducted with fish models, which may range from life-like models made from marine ply, as used by GBRMPA for coral trout (1979), and by Samoilys and Carlos (1992) for serranids (rounded tail models) and acanthurids (forked tail models, Figure 3.4a), to PVC pipe cut into lengths as used by Bell et al (1985). Both types are discussed in the AIMS manual (English et al 1994). Clearly, the more life-like the models the better. Samoilys and Carlos (1992) used a simplified model made of marine ply in Fiji (Figure 3.4b).

Figure 3.4 (a) Australian and (b) Fijian Fisheries officers training with plywood fish models.

Figure 3.3 Accumulation of numbers of fish over time during the progression of 10 minute point counts, for sedentary acanthurids (modified from Samoilys and Carlos 1992).

1. Construct a set of fish models of sizes ranging from the smallest lengths included in the visual surveys to the largest fish normally encountered, i.e.: from 11cm to 100cm. Increments of 1cm to 2cm are recommended to allow for even length estimation training over the whole spectrum of sizes likely to be encountered in the field. Models are strung end to end along thin ropes and should hang vertically in the water. The ropes are anchored in shallow water where trainee observers record their estimates of lengths on snorkel.

It is recommended that the whole spectrum of sizes is included in the set of models (cf. Bell et al 1985, English et al 1994) because the aim of the training exercise is to train observers to estimate fish of all sizes equally well (Samoilys and Carlos 1992). The set recommended by English et al (1994) based on Bell et al (1985) approximates the normal size distribution of a population of fish with mean size of 50cm. This results in biased training with more practice on the mid-range sizes and less on the small and large fish.

2. Each trial involves a sub-set of 50 models selected randomly from the whole set. The actual length is marked on the back of each fish. Trainee observers swim along the line of models at a distance of 2-3m from the fish models, recording their estimated lengths with pencil and slate. They then compare their estimates with the actual lengths. Paired t-tests are useful tests for this comparison. In general, length training using wooden fish involves around six trials (Samoilys and Carlos 1992) before observers achieve acceptable accuracy.

Prior to starting the trials observers may "key in" to a couple of models of known lengths.

3. Trials are continued until there is no significant difference between the estimated lengths and the actual lengths. A graphic illustration of the results may be plotted to demonstrate an observer's bias. Estimated lengths are plotted against actual lengths (Figure 3.5). The solid line, where y=x, represents perfect accuracy. Thus points below the line indicate the observer is under-estimating sizes, and points above the line indicate the observer is over-estimating sizes. If the points are widely scattered both above and below the line it suggests the observer is inconsistent. If this is the case they should not be used for fish counts.

4. Retraining of observers is necessary if they have not been engaged in visual census work for a long period (e.g. over 4 months).

Figure 3.5 Fish length estimations from plywood models by an experienced observer (redrawn from Samoilys and Carlos 1992).

3.4 Field procedures

Important considerations:

Always standardise the procedure for each census

Try to simulate an instantaneous "snap-shot" count

Remember to swim slowly

Note: a field trip equipment check list is provided at the back of the manual.

The following procedures are based on the example of sampling a reef slope habitat.

1. Select at least two study sites each 1km in length along the reef slope (i.e. parallel to the reef crest). Technically speaking these sites should be selected *randomly*. In reality they are often selected *haphazardly*. Random selection can be done using boat travel time. For example the reef slope may take 30 mins to travel along by skiff. Choose random numbers between 0 and 30 to represent minutes of travel to the start of the next site, ensuring sites are separated by at least 100-200m. Sites extend to 15m in depth, the limit imposed by repetitive SCUBA diving.

2. Select the locations of the replicate visual census counts within each site, randomly, in terms of site length, width and depth. For example a 1km site takes 5 minutes (300 seconds) to travel along by skiff, therefore choose census locations from random numbers between 0 and 300 without overlapping any locations. If the site is narrow (i.e. shelves steeply) counts are located in a line. If the site is wide because the reef slope is gradual, counts are located at varying distances from the reef edge - again locate these distances randomly or haphazardly. Transects are laid parallel to the reef edge or crest. Choose 10 - 12 locations (replicates) per site prior to starting the survey, and then vary the order in which the replicates are done.

Replicate sampling units, i.e. UVC counts, are placed *randomly* within each site to ensure each census is *independent* of any other census. Again, many surveys involve haphazard locating of replicates rather than random, because it is quicker. If this is done it is critical that bias is minimised. For example "good spots" are not chosen to give high counts of fish density! The order in which replicates are done along the reef should be randomised to avoid any bias associated with fish movements.

3.5 Fish counting techniques

1. First, count the larger mobile species. e.g. roving serranids such as *Plectropomus* spp.; lethrinids, larger lutjanids such as job fish (*Aprion viriscens*) and bass (*Lutjanus bohar*), large scarids, the mobile acanthurids (*Naso* spp.), etc. This ensures these types of fish are counted before they leave the census area.

2. Second, concentrate on the smaller more sedentary fish such as the smaller lutjanids, other scarids and acanthurids. These types of fish are less likely to leave the census area because they are less mobile.

3. Don't count any fish that enter the census area after the stop watch has started (= *incoming fish*).

Samoilys and Carlos (1992) demonstrated that overestimating numbers of fish in a census is a significant problem if **"*incoming*"** fish are not distinguished. These are fish that enter the census area after the census has started. In conducting a visual census the observer attempts to simulate an instantaneous or "snap-shot" count, which "captures" or counts only those fish that are in the census area at t_0 (time=zero), the start of the count. Any fish that enter the area after t_0 should be disregarded because they will inflate the density of fish in the census area. With practice, this is not difficult to do - simply ignore those individuals that cross the census boundaries into the count area. Underestimating numbers of fish is a similar but opposite problem, which has long been recognised and is usually attributed to the observer simply not noticing some fish (Sale and Sharp 1983). Underestimating may also be caused by missing individuals that leave the census area after the count has started but before the observer has counted them. This is not, however, offset by incoming fish - there is no logical reason as to why they should be equal. Different strategies are required to minimise both types of error. In the case of overestimating, *incoming* fish must be distinguished, and they are not included in the count. In the case of underestimating, mobile species are counted first and the number of species included in a count is kept to a minimum. These issues were discussed at length during a

workshop held during Phase 1 of the ACIAR/DPI UVC project (see proceedings: Samoilys 1992).

4. A number of physical parameters should be recorded for each census such as weather (cloud cover, sea state), time of day, tide, depth (minimum and maximum) and water visibility (see sample datasheet).

These parameters are easily recorded and provide measures of variables that may potentially affect fish densities. For example high rainfall and low water visibility may create difficulties in counting fish, giving unexpectedly low densities - this can be quantified if weather and visibility have been recorded. Refer to the AIMS manual for further details on recording environmental parameters (English et al 1994: pp. 7-11). Selecting parameters to measure relates to the objectives of the study. If the general "health" of the reef is to be assessed, several environmental parameters (e.g. coral cover) should be measured, as described in the AIMS manual.

3.5.1 Stationary point counts

This section describes the procedures for conducting a census of 7m radius area based on the method of Bohnsack and Bannerot (1986) modified by Samoilys and Carlos (1992).

SCUBA divers:

Observer One who counts the fish

Observer Two who acts as "dive buddy" and collects the substrate data.

It is preferable to maintain the same observer as Observer One throughout the surveys to reduce errors associated with observer differences.

Equipment:

(1) **Observer One**: - Pencil and slate with pre-determined species list prepared on waterproof datasheet with all other variables to be recorded listed e.g. date, time, tide, depth, visibility etc. (see example datasheet). Stop watch.

(2) **Observer Two**: - Tape-measure.

The following procedures are based on the example of sampling a reef slope habitat.

1. Anchor the boat and swim a fixed distance (e.g. 20 fin beats), from the boat along the reef edge before starting the count. Observer One should lead and Observer Two follow.

The census is started away from the boat because boat noise, anchoring and divers entering the water may have disturbed the fish.

2. After the last fin beat dive down slowly towards the reef bottom, Observer One leading and Observer Two following. As soon as the reef bottom is visible and/or fish can be seen over it, start the stop watch and begin counting and recording the fish. At the same time visually fix a central point on the bottom. This marks the centre of the circular point count. Simultaneously estimate the radius of the count area - e.g. 7m from the central point. Note features of the habitat to mark the circular boundary of the census. Continue to swim slowly down to the central point, depending on how many fish are visible from above. If there are many large visible fish remain up in the water column for longer.

3. When close (approx. 3m) to the bottom Observer One indicates the centre point to Observer Two. Observer Two then drops slowly to the central point and remains there, stationary, until Observer One has finished counting. Observer One continues counting, turning slowly to search a 360° circle (Figure 3.6). Observer One then swims around the area to search for smaller, cryptic fish.

Figure 3.6 Diver (Observer One) recording fish on the Great Barrier Reef from the centre of a stationary point count.

4. At the end of the fixed time of the count (e.g. 7 minutes), Observer One takes the tape-measure from Observer Two to measure the count radius that was estimated visually. To do this, Observer One swims to a point on the boundary, attaches the tape-measure to the substrate and swims across the census area, past Observer Two, in a straight line to the other side of the circle, i.e. two radii. The mean radius is used to calculate the actual area of the census.

5. Observer Two follows Observer One in step 4 above and records the substrate. Observer Two measures the distance of habitat categories beneath the tape using the *line intercept method* (UNESCO 1984, English et al 1994: pp. 34-51). Examples of habitat categories are: live (hard) coral, dead coral, sand, rubble, algae, soft coral.

Substrate measurements can be done at various levels of detail depending on the information required. For example coral life form categories may be identified to provide a morphological description of the reef community. Coral species may also be identified. The AIMS manual (English et al 1994) should be consulted for further details. For fisheries stock assessment purposes, substrate measures are recommended during visual surveys of fish populations to provide a broadscale, but quantified, assessment of the reef habitat of the study areas. This is useful for detecting or monitoring degraded habitat such as coral destruction caused by dynamite fishing and coral die-off caused by siltation from river run-off.

6. Observer One estimates water visibility when winding up the tape-measure across the diameter of the count. Look ahead along the tape to the point of attachment (0m on tape). Wind up the tape until this point becomes visible, note this distance on the tape-measure. The visibility may be greater than the diameter of the point count. If so it is recorded as, for example, >14m.

Water visibility is one of several environmental parameters that may be measured to characterise the study area and the conditions at the sites when censusing.

3.5.2 Strip transects

This section describes the procedures for conducting a 50 x 5m area transect census.

SCUBA Divers:

Observer One who counts the fish

*Observer Two* who acts as "dive buddy" and collects the substrate data

It is preferable to maintain the same observer as Observer One throughout the surveys to reduce errors associated with observer differences.

Equipment:

(1) **Observer One**: - Pencil and slate with pre-determined species list prepared on waterproof datasheet with all other variables to be recorded listed e.g. date, time, tide, depth, visibility etc. (see example datasheet).

(2) **Observer Two**: - Tape-measure and stop-watch.

A 5m length of 3mm buoyant rope is tied between the two observers to mark the transect width. The rope is most easily attached at the divers' elbows. Half way along the rope attach a small net float - this helps to keep the rope up in the water column so that it doesn't snag on the coral (see Figure 3.7). The transect width rope ensures the observer is aware of the transect boundary.
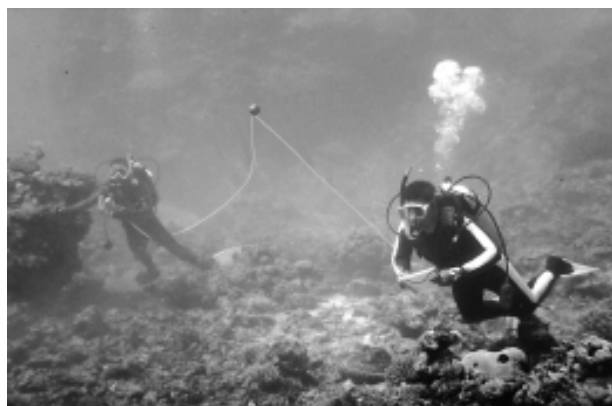
Figure 3.7 Divers conducting a strip transect in Fiji: note the connecting rope with float.

1. Having anchored the boat, the two divers swim to the bottom. They then attach the 5m connecting rope (transect width marker) between them, and swim apart to extend the rope to the transect width. They then swim a fixed number of fin beats (e.g. 20) away from the boat along (parallel to) the reef slope before starting the transect.

The census is started away from the boat because boat noise, anchoring and divers entering the water may have disturbed the fish.

2. Observer Two attaches the tape-measure to the substrate, indicates the start of the count by pulling on the connecting rope to Observer One and starts the stopwatch. Observer Two maintains a constant swim speed (e.g. ~6m/min) and lays out the tape in a straight line (Figure 3.6) as Observer One records fish within the transect area. Observer One visually projects the boundaries of the transect ahead; the distance ahead depends on water clarity. Observer One zigzags across the transect to search the area thoroughly.

Figure 3.8 Diver (Observer Two) laying the tape and maintaining constant swim speed in a strip transect.

The procedure described here involves simultaneously laying the transect tape-measure and counting the fish, unlike the method described by English et al (1994) where the tape measure is laid first. The simultaneous procedure adopted here (also used by Fowler 1987) is highly recommended because it avoids problems of fish disturbance caused by laying the transect tape. Mapstone and Ayling (1993) also recommend the simultaneous procedure, though they prefer to estimate the transect width and then measure the estimate rather than use a connecting rope. The technique suggested here involves the two divers swimming more or less parallel because they are attached by the connecting 5m rope. This improves the accuracy of the observer's visual projection of the transect boundaries ahead.

As described above in Fish counting techniques, Observer One concentrates on the larger more mobile species first within each visible section of the transect ahead, and then counts the smaller more sedentary species at closer range.

3. When Observer Two reaches the 50m end of the tape-measure s/he signals the end of the count to Observer One by pulling on the connecting rope. This point should coincide with the fixed time of the transect, e.g. 7.6 minutes for 50m x 5m transect at ~6m/min.

The ACIAR/DPI UVC Project determined the optimal speed for 50m x 5m transects was $33m^{-2}$ min^{-1} (Samoilys and Carlos 1992).

4. See Step 5 of the point count method for substrate recording.

5. See Step 6 of the point count method for water visibility recording.

3.6 Observer bias

Differences in the ability to count fish and estimate fish lengths will occur between observers. UVC requires training both in terms of identifying species, being proficient in estimating their abundance, and in being able to accurately estimate their lengths. The latter can be tested, as described in the section above. Accuracy in estimating abundance is difficult to test, because the actual or real number of fish is not known. However, observers can be compared. This is useful when a new observer is being used, or if two or more observers are required for a particular project. In such situations the observers should conduct a set of counts in the same area and compare their estimates (see Samoilys and Carlos 1992 pp.52-53).

3.7 Calculating density and biomass

This section briefly outlines the procedures for calculating density and biomass from raw UVC data. Full details on processing and analysing data are given in Chapters 5 and 6.

UVC Surveys

3.7.1 Calculating density

1. Density = n/y

where *n* = number of fish (individuals) of species *a*

and *y* = census area

This calculation is done for every species (*a,b,c,* etc) in each count or replicate (= one record in the database, see Chapter 5).

In the 7m radius point count method described above the census area, $y = 154m^2$.

Example of data:

replicate 1:   10 *Acanthurus lineatus* in $154m^2$

replicate 2:   13 *Acanthurus lineatus* in $154m^2$

etc. to replicate 12

If all census areas are the same each replicate count represents a density value per unit area of $154m^2$. In the method described in this chapter the point count census area is estimated visually and then measured with a tape-measure. Thus, each replicate count may have a different census area, *y*. With transects the census area is the same for each replicate i.e. $250m^2$.

2. Standardise each replicate or record to a unit area so that replicate counts are comparable. Typically $1000m^2$ is used as a standard area.

Thus: Density = $\dfrac{n \times 1000}{y}$

Chapter 5 describes procedures for doing this within the Access database.

3. Preliminary analyses involve the calculation of mean density + standard error per site for each species as described in Chapter 6.

3.7.2 Calculating biomass

1. Convert fish lengths to weight or biomass using length-weight relationships, having first standardised the recorded fish length (see Chapter 5).

The relationship between total length and fish weight is defined by the ***length-weight relationship*** of the form:

$wt = aL^b$

where wt = weight

L = fish length

a and b are constants

Sparre and Venema (1992) describe procedures for determining length-weight relationships. It is preferable to use length-weight relationships that have been obtained from the study area (e.g. from creel surveys, see Chapter 4), but in practice they are not always available. The constants *a* and *b* have been calculated for a wide range of coral reef fish species. The following publication provides a wide range of length-weight relationships for fish exploited in the Pacific, and was used in the ACIAR/DPI UVC Project (Samoilys et al 1995, see also Chapter 5):

Kulbicki M, Mou Tham G, Thollot P & Wantiez L (1993) Length-weight relationships of fish from the lagoon of New Caledonia Naga 16 (2-3): 26-29.

Other publications of relevance to the Pacific Islands are:

Wright A and Richards AH (1985) A multispecies fishery associated with coral reefs in the Tigak Islands, Papua New Guinea. Asian Marine Biology 2: 69-84.

Loubens G (1980) Biologie quelques especes de Poissons du lagon neo-caledonien. Cahiers de l'Indo-Pacifique 2: 101-153.

Where a particular species is not represented in any of the published length-weight relationships, the closest species based on genus, body shape, and maximum length is selected.

Each data record consists of an estimated length for each individual fish. Estimations are either Total Length, TL, for fish with rounded tails e.g. Serranidae, or Fork Length, FL, for fish with forked tails e.g. Acanthuridae (see section 3.3 above); both are usually recorded in centimetres.

The published length-weight relationships may use mm (e.g. Wright & Richards 1985), but the UVC estimates are in cm. Therefore they must be converted to mm (x 10) first. Similarly the published length-weight relationships may be as Standard Lengths, SL (e.g. Loubens 1980), whereas the UVC estimates are TL or FL. Therefore, the UVC estimates must be converted to SL first, using the equation provided in the publication.

2. Example

Record: - *Lethrinus harak* **20cmFL**

Length-weight relationship for *Lethrinus harak* **in cm from Kulbicki et al (1993):**

wt = aFL^b

a = 1.54 x 10^{-2}

b = 3.043

wt = 140.1g

3. As in section 3.7.1 above, fish weight per unit area is calculated to give standardised biomass estimates for further analyses (see Chapter 5 for details).

CHAPTER FOUR: FISHERY (CPUE) SURVEYS

David Die

Some definitions:

Fishery survey: study to sample characteristics of the fishery (fishing effort, gear types) or of the fishery catch.

Scientific survey: study to sample fished stocks on board a scientific vessel with commercial or scientific gear.

4.1 Why do surveys?

Given the size of most fish populations harvested by humans the collection of information on such populations is an expensive and laborious task. There are two sources of data that can be collected from these populations: the fishery itself and scientific surveys. Data from the fishery is generally cheaper to obtain and therefore can be collected in large numbers. However the source of this data is limited to the times and areas where the fleet operates, it may therefore not accurately represent the real fish population. Data from scientific surveys can be more accurate than fishery data, however, the collection costs are so much higher than these data are often a lot less precise than fishery data. In fact for most major world fisheries, the fishery is the main data source for stock assessments. It is essential, however, to use scientific surveys to ensure that fisheries data are accurate and do not present a biased view of the status of fished stocks.

Fishery surveys also provide information on the operations of the fishery (gears types, fishing patterns, fishing grounds) which are essential in understanding the impact of fishing upon the stock. Unless we can measure the amount of fishing on a given population we will not be able to relate changes in population abundance with the impacts of fishing. For instance even if we could precisely census (count all individual fish) a fish population and describe changes in abundance with time and area we would still need to know something about the amount of fishing before we could relate changes in abundance with changes in fishing pressure.

All fishers target certain species and sizes of fish. Even the less selective gear types will always be more effective at catching certain species/sizes. Therefore the fishery catch will always represent a "biased" sample of the fish community present. Most scientific sampling methods are similar to fishing gear, they will collect certain species/sizes

preferentially. Like fishing gear, most sampling methods can only be operated in certain places/times. Properly designed experiments, however, are often used to quantify the biases associated with scientific sampling. Therefore it is easier to obtain unbiased samples from scientific surveys than from fishery surveys. There is an extensive literature on the design of fishery surveys. For general texts see Bazigos (1974), Brander (1975), Caddy and Bazigos (1985), or more recently Sparre and Venema (1992). For design of scientific surveys see Saville (1977) and for a review of statistical models as applied to analysis and designs of surveys see Doubleday and Rivard (1983).

4.2 Using survey data in stock assessments

To combine data from both fishery and scientific surveys we have to establish how the variables measured (abundance, catch rate, sampling effort, fishing effort) relate to one another because of the differences in the characteristics of the two types of surveys, e.g. in: species/size selectivity, sampling coverage, sample sizes, sampling design, accuracy and precision. Only then can we use the combined data in an assessment model of the fishery.

Most fish stock assessments rely on data collected from both scientific surveys and fishery surveys. Fishery models make use of both sources of data in order to develop an assessment of the status of the stock. Most of these models, however, rely on very strong assumptions about the relationships between the variables (abundance, catch rate, sampling effort, fishing effort) measured from the two types of surveys.

Fishery surveys are not different to any other population sampling: samples must be taken to best represent the population variable to be estimated. Fished stocks are a special case of biological populations and they tend to be defined by a mixture of biological (e.g. group of individuals which share unique breeding locations and times) and operational (group of populations that are fished by the same fleet or managed as a distinct unit) attributes. Fishery surveys, however, will never allow us to sample the entire stock but rather the part of the stock that is caught by the fishing fleet - the catch.

The most important information we can obtain from a fishery survey is the annual catch and annual fishing effort.

Often we can only estimate either the total catch or the total effort from a fishery survey but not both. However, if given an estimate of catch per unit of fishing effort (*CPUE*) and a measure of one parameter (e.g. catch) we can calculate the other parameter (e.g. effort). Unfortunately it is not uncommon to have a fishery for which we only know CPUE. CPUE alone will not tell us anything about the impact of fishing or the potential catch that can be taken from a stock, unless is monitored over a long period of time.

4.3 Types of fishery surveys

According to how and where they are conducted, there are several types of fishery surveys (e.g. onboard, questionnaire, creel, frame). Onboard surveys are conducted while the vessel is fishing, and therefore allow for the precise estimation of time and location of catches, as well as the opportunity to describe discarding practices if they exist. These surveys, however, are very time-consuming and the amount of data collected are limited to the number of observers at the time.

Frame surveys are conducted to establish the optimal design for a fishery survey and are conducted before a major survey begins. Their main objective is to collect enough baseline information for selecting sampling sites, sampling frequency, and appropriate sampling techniques.

Fish are sometimes landed in places that are not easily accessible or are landed in many different places at unpredictable times. This is a common characteristic of many artisanal, subsistence and recreational fisheries. In such cases fishers have to be interviewed at a time other than the time at which they land their fish and in a place that may be far from their landing area. Such surveys are known as questionnaire surveys and they rely on the knowledge of the fishery and the fishery operation held by the interviewee. They have the advantage that information on the operation of the fishery and the economics of fishing are easier to collect because the fisher has more time to answer questions. This contrasts with creel surveys, those conducted at landing stations, where the selling or processing of the catch is the first priority of the fisher.

Note: a field trip equipment checklist is provided at the back of the manual.

4.4 Creel surveys

Fishery surveys conducted at the landing place are known as creel surveys. Landings tend to be restricted in time and place to a few ports and times of day, and therefore creel surveys allow for the sampling of large quantities of fish. Creel surveys generally produce the most comprehensive source of data on a fishery.

Creel surveys are fishery surveys conducted in the place and time of landing. Apart from logbooks, they are the most efficient method for collecting comprehensive information on catch and fishing effort. They also allow for the collection of biological samples that can be taken to the laboratory for further analysis.

Creel surveys can be designed to estimate many different things: total catch, total fishing effort, CPUE, species composition of the catch, length frequency of fish in the catch, gear numbers and gear types etc. Commonly, a given creel survey will collect all this information for a particular area and time, but the information will be pooled with other creel surveys to produce an overall description of the fishery. The experimental design of creel surveys is very important and should be related to the specific objectives of the study (see Chapter 2). A creel survey designed to estimate the total catch of a stock should be designed such that those sampling units (ports, boats, times) where the majority of the catch is landed are sampled the most. By comparison, a creel survey aimed at cataloguing the species composition of the catch will put more sampling effort in those sampling units where the species diversity is the highest. This is called optimal sampling, where sampling effort is apportioned (or weighted) in relation to the proportion each stratum contributes to the variability of all strata (see Chapter 3, section 3.2), where strata here refer to ports, boats or times.

In the design of creel surveys we should follow the general rules of sampling design (see Andrew and Mapstone 1987 for a review). The accuracy and precision (Chapter 2) of survey estimates are closely related to the design of the survey and the characteristics of the fishery. It is also important to establish which statistical tests are appropriate for interpreting the results (see Chapters 2 and 6).

CPUE surveys

The study by Samoilys et al (1995) is used here to illustrate an example of a creel survey conducted to estimate catch per unit of effort within a study area. The second objective of the study was to estimate the species composition of the catch, and the third objective was to describe the fishing operation (gear, boat types, crew). To achieve their primary objective Samoilys et al focussed their creel surveys on those landing areas which were used by fishers operating in the study area. Those landing areas were determined in a preliminary frame survey.

The creel survey form should be designed to fulfill the objectives of the study only. Including extra information is not recommended because it will tend to take time that could have been used in sampling other landing units.

There are many examples of creel survey forms in the literature (e.g. Brander 1975), and if possible it is best to use or modify a form that has proven to be well designed, rather than try to design a new form. The form used by Samoilys et al (1995) is shown in Table 4.1.

It is recommended that the appropriateness of the survey form be tested before the final surveys are conducted, for example during the frame survey.

Testing of the form must determine a range of things and some general rules can be applied to both creel and questionnaire forms. Examples of these rules are: determine whether all questions are easily understood by both interviewer and interviewee; determine whether all answers can be assigned to particular categories; determine whether answers are given in the same units (e.g. fish weights); determine whether there is enough space to write all the information. Once tested and corrected the form should not be modified for the duration of the study.

It is essential that all questions in a creel survey form are filled out, and that the information recorded conforms to the same standard established for all persons participating in the survey. It is a good habit to tick all questions in the survey forms to confirm they have been asked, and thereby ensure the unequivocal transcription of survey results.

4.5 Questionnaire surveys

If it is not possible to get fishery data at the landing place, a questionnaire survey can be conducted within the fishing communities, companies and processors.

Questionnaire surveys are based on interviewing members of the public (households, individuals) that are potentially engaged in fishing activities.

A questionnaire survey will produce less reliable catch data, because the catch can not be measured, counted or classified. The quality of the information will depend on the memory of the interviewee and his/her willingness to provide it. Questionnaire surveys, however are more effective than creel surveys at providing summary information on the operational characteristics of the fishery. Used appropriately they can also provide rough - but very valuable - estimates of catch, CPUE and fishing effort.

The design of questionnaire surveys and of questionnaire forms follows the same considerations outlined above for creel surveys.

Rawlinson et al (1995) provide a detailed discussion on questionnaire design and on the logistics of conducting questionnaire surveys for assessing subsistence and artisanal fisheries in Fiji.

The study by Samoilys et al (1995) is used here to illustrate an example of a questionnaire survey in which the principal objective was to obtain estimates of the total number of units (boats, people) participating in subsistence and artisanal fisheries that operated within the sample areas where UVC surveys were conducted. A secondary objective was to estimate CPUE and the operational characteristics of the fishery (seasonal effort patterns, gear types). Samoilys et al (1995) conducted the questionnaire surveys in those villages which were identified in the frame survey as most likely to host fishers operating in the study areas.

The information obtained in questionnaire surveys consists of a series of answers to a list of questions. The questionnaire survey form used by Samoilys et al (1995) is shown in Table 4.2. Due to the great variety of potential answers to any question it is essential that questionnaire survey forms try to classify the answers into categories. It is also essential that these categories are clearly mutually exclusive such that interviewers cannot make subjective choices. It is important that interviewers make the effort of categorising the answers during the interview rather than during the transcription of data to a database. If an interviewee answers a question in such a way that the interviewer cannot decide what category the answer belongs

Table 4.1. Creel survey form used by Samoilys et al (1995) to survey subsistence and artisanal fisheries in Solomon Islands and Fiji.

CREEL SURVEY

DATE: TIME: LANDING SITE:

RECORDER: FISHER: CREW: LICENSE/BOAT:

Have you captured any of these fish in the study area?   (yes/no)

How often have you fished in the study area over the last month?

AREA	A	B	C	D	E	F	Total
DAYS							

What gear did you use?

Spearfishing		Gillnet		Handline		Trolling	

DETAILS OF LANDING CATCH

SPECIES	AREA														
	Total		A		B		C		D		E		F	
	number (#)	weight (w)	#	w	#	w	#	w	#	w	#	w	#	w

EFFORT (Hours)

	Total	A	B	C	D	E	F
All gears							
Spearfishing							
Gillnet							
Handline							
Trolling							

CPUE surveys

to, s/he should ask the interviewee to expand on the answer so that s/he can decide on the category. All questions in the form should be asked and it is recommended that a tick is placed in the survey form to confirm this.

4.6 The analysis of survey information

The study of Samoilys et al (1995) provides a useful illustration of the analysis of survey data. The main purpose of both the creel and questionnaire surveys in their study was to estimate the amount of fishing within their study areas. This is achieved by following the steps described below. Further details and examples on the analysis of fishery survey data are provided in Chapter 6.

It is essential to start by describing the main characteristics of fishing activities within and outside the study area. The first step in the analysis should be describing the frequency of usage (proportion of trips sampled) in each area sampled. This should be done for both the questionnaire data and the creel survey data. The next step is to describe the type of gear used and the species composition of the catches in each area sampled, also from both sources of data.

Once the basic characteristics of the fishery have been defined we can estimate the catch, effort and catch per unit of effort, $CPUE_s$ for each sample area, s. $CPUE_s$ in weight and numbers should be estimated for each gear type and each species-group from both the questionnaire and the creel surveys.

The appropriate unit of fishing effort should be investigated by looking at the distribution of length of trip (in hours), for each gear type. If this distribution is not too variable (e.g. 90% of all trips are within plus or minus one standard error of the mean) then hours fished can probably be ignored and the effort unit should be the fishing trip.

It is possible that several gears are used in the same trip. If this occurs on only a few occasions the trip should be assigned to the gear that caught most of the fish. If using more than one gear type is common, the fishery may have to be defined as a multiple-gear fishery, and fishing effort should be calculated for the mixture of all gears. For example in the study by Samoilys et al (1995) three main gear types were found in the Solomon Islands: handline,

gillnet and spearfishing, from either paddle canoes or outboard-powered canoes.

If the structure (number of crew, size of boats, types of gears used) of the fishing fleet is very variable, it may be necessary to break the fishery fleet into categories. Fishing effort and catch should then be calculated for each category, and standardization factors should be estimated in order to combine fishing effort across fleet categories (Robson 1966). Standardization involves an analysis of variance to determine differences in fishing power between different categories (e.g. fishing gears/vessels). Correction factors are used if categories have significantly different fishing power at 5% (p<0.05).

Standardising fishing effort is done by using a simple analysis of variance (ANOVA, see Chapter 6). First, estimate the log CPUE for each sample (record) in each fleet category. Second, group these observations according to time-area strata (to ensure we compare CPUE of vessels fishing the same population). Third, perform a two-way ANOVA with fleet category and time-area as the two factors. The coefficients obtained by the ANOVA model for each fleet can be used as fishing power factors (for a review see Robson 1966). For example Samoilys et al (1995) categorised the creel and questionnaire survey data obtained from the fishery in Marovo Lagoon, Solomon Islands, into the following categories: (1) time of day (dawn, day, dusk, night); (2) boat type (paddle canoe or outboard-powered canoe); (3) gear type (handline, spear, gillnet). The CPUE data were further stratified by time (surveys), fishing area and species group (carnivores/ herbivores). Fishing power analyses, using ANOVA, compared fishing effort from the different gear/vessel combinations, and found no differences except between spear and handline gears, for carnivores from paddle canoes in creel data. Spear fishing had lower fishing power. However, since this difference was only detected in one combination, it was decided not to apply correction factors when estimating relative fishing effort across the fishery. Handline was the dominant fishing method and therefore, further analyses focussed on data for handline fishing only (chapter 8 in Samoilys et al 1995).

The following sections outline procedures for calculating fishing effort and catch. Full details for processing and analysing data are given in Chapters 5 and 6.

Table 4.2. Questionnaire survey form used by Samoilys et al (1995) to survey subsistence and artisanal fisheries in Solomon Islands and Fiji.

<div align="center">

QUESTIONNAIRE SURVEY

SECTION 1. RESPONDENTS IDENTIFICATION

</div>

INTERVIEWER		INTERVIEWEE	
DATE		OCCUPATION	
AREA		AGE	
VILLAGE			

<div align="center">

SECTION 2. FISHING OPERATION

</div>

Main use of fish caught	Self consumed		sale		other	
Frequency of fishing	daily 3-7 times week		weekly 1-2 times a week		monthly < 1 week	
Length of fishing trip	day		2-3 days		more	
Fishing status	indiv. fisher		crew in boat		owner/skipper	

<div align="center">

SECTION 3. FISHING AREA

</div>

Do you fish inside or outside of the sample areas? How often (rank), when (tick), how (what gear) and what species do you target the most?

		Month											Gear	Target species
AREA	Rank	J	F	M	A	M	J	J	A	S	O	N	D	Gear	Target species
A															
B															
C															
D															
E															
F															
Outside															

Notes:

Table 4.2. continued.

QUESTIONNAIRE SURVEY

SECTION 4. CATCH

Can you describe your most recent trip?

Date of trip: Length of trip (days/hours)   Boat type   Crew

AREAS FISHED	A	B	C	D	E	F	Outside
Rank them							
Gears used							

Species	Number	Weight	Gear used

Can you describe your best trip/catch during this last year

Date of trip (approx):   Length of trip (days/hours)   Boat type   Crew

AREAS FISHED	A	B	C	D	E	F	Outside
Rank them							
Gears used							

Species	Number	Weight	Gear used

Notes:

4.6.1 Estimation of annual fishing effort

Annual fishing effort can be estimated from data obtained in questionnaire and creel surveys. For both data sets the fishing year should be divided into seasons according to the seasonal fishing pattern identified in the data. If there is no seasonal pattern the weekly pattern should be identified. The relative proportion of effort $f_{d,s}$, in each period of the week, d, and in each fishing season, s, should be determined from the questionnaire survey and creel survey data. The number of seasons (e.g. dry and wet) and the number of weekly periods (e.g. mid-week and week-end, or Monday to Thursday, Fri, Sat, Sun) will have to be determined by analysing the data. It is possible that no seasonal or weekly pattern is found, in such cases it will be assumed that all fishing days are equivalent ($f_{d,s}$ = *constant*).

Let's now assume that $f_{d,s}$ = *constant*, using the example of Samoilys et al (1995). The questionnaire data provides the proportion Pq_s of interviewees that fished in each UVC study area, s:

Pq_s = Number of questionnaires that fished in area s / total number of interviews

Given census data on the populations of each village/town or an estimate of the proportion of households/persons interviewed during the surveys it is possible to estimate the total number of persons participating in the fishery. The estimate of annual effort in each area, F_s, is then obtained as:

F_s = Pq_s x Number of persons in the fishery x average number of trips/year/person

where the average number of trips/year/person is directly estimated from one of the questions of the questionnaire form (frequency of fishing).

4.6.2 Estimation of annual catch

The estimation of annual catch, *Cs*, will be done from the $CPUE_s$ (in weight or numbers) and the annual effort:

C_s = F_s x $CPUE_s$

Estimates of annual catch, effort and catch per unit of effort are calculated for each fleet category.

CHAPTER FIVE: DATA STORAGE AND MANIPULATION

Gary Carlos and Robert Koelldorfer

5.1 Introduction

A database allows large amounts of data to be organised and stored. A well-designed database system can represent a comprehensive history of a fishery, in which changes in the fishery can be monitored, analysed, interpreted and future comparisons made. This chapter describes how to set up and manage a database for storing and processing data collected from underwater visual census (UVC) and fishery (CPUE) surveys (see Chapters 3 and 4). There are a number of steps involved in setting up a database system: design and development of the database; designing a standard operating procedure for using and managing the data and setting standards and procedures to ensure data accuracy and reliability (Bainbridge and Baker 1994). The chapter by Bainbridge and Baker on "Database Design and Operation" in the AIMS manual (English et al 1994) provides a very useful and clearly written synthesis, and is strongly recommended for learning the principles, concepts and step by step process of database design and management. This chapter draws on the work of Bainbridge and Baker and specifically addresses the database system used by the ACIAR/DPI UVC Project (Samoilys et al 1995).

There are essentially two types of software applications that can be used to store and maintain an organised collection of data: a true database such as Access, or a spreadsheet such as Excel. There is often debate on the relative merits of databases versus spreadsheets. Although the flexibility and ease of use of spreadsheets is tempting, they are inappropriate for large data sets because of problems in data consistency and integration (Bainbridge and Baker 1994). These authors discuss clearly the advantages of databases and the potential problems of spreadsheets. There are a large number of factors involved: data consistency, data efficiency, data quality, data analysis, data integration, speed, data extraction, ability to program and storage methods. Basically, a database should be used for storing data and a spreadsheet for working on sub-sections of the data. The advantage of a spreadsheet such as Excel lies in its ability to summarise, manipulate and use graphical and basic statistical features on sub-sections of data.

A database consists of tables which contain *fields* and *records*. The fields are columns and represent different attributes of the object or event that is being recorded (such as lengths and numbers of fish). The records are rows, and each record represents a different set of observations about the object or event. Bainbridge and Baker (1994) clarify these terms, explain the difference between *relational* databases and *flatfile* databases and define optimal procedures for designing a database. The following summarises some of their main points.

A *relational* database, such as Access offers more efficiency by splitting the data across a number of tables which are related to each other by a *linking field*. Thus tables share a common field which identifies which records are to be linked. For example, in the UVC database there are two tables with specific details on (i) the replicate - one visual census and its physical conditions (e.g. date, time, observer, depth and substrate), and (ii) fish identification (e.g. species, number and lengths of fish). A field named *Sample ID* is common to both tables and thereby links them to each other. The effectiveness of such a relational database is dependent upon the user's knowledge of how information in the tables is related.

The advantages of a relational database can be summarised as follows:

a) A set structure to which the data must conform

b) No set limit in the number of records

c) Efficient in storage space and CPU speed - the duplication of data is reduced

d) The ability to add data validation conditions and checking programs to minimise errors in data entry

e) The ability to retrieve or extract data using complex in-built queries

f) The foundation for integrating different data sets into regional and international data sets

g) In-built programming languages and basic statistical routines.

This chapter of the manual provides a general guide to operating a relational database. All relational databases share certain characteristics. Although Microsoft Access 2.0 is the specific format used in these examples (due to its wide availability among South Pacific fisheries

organisations), the logical steps that are specified are appropriate for any database program. Examples of storing, manipulating and analysing UVC and CPUE data for coral reef fish stock assessment are given, using real data from sampling surveys conducted in Fiji and Solomon Islands (Samoilys et al 1995).

5.2   Building the database

The design of a database will evolve naturally in accordance with the sampling design of a project. Designing a database requires careful planning and the final design is usually the result of a number of modifications. A well-designed relational database should have the following aspects:

a) Familiarisation with the data being collected

b) Well-designed data sheets

c) Arrangement of information into groups of data

d) Database tables which reflect the data sheets and the groups of data

e) Definition and validation conditions for each field

f) Careful identification of replicates

g) Inclusion of any variables required for manipulation (e.g. date, time)

h) Adequate data checking procedures

i) Testing procedures to ensure the database reflects the data being collected.

A database management system should ensure that the data are defined, described and entered correctly, and are backed up. Incorporating a documented *Standard Operational Procedure (SOP)* is strongly recommended. A SOP should detail all the procedures for operating the database, the methods used for data checking, a list of any codes used, instructions for how to backup and archive the data and responsibilities for data handling (Bainbridge and Baker 1994).

5.2.1 Procedures for creating a database

This section gives general steps which must be followed in creating a well designed database and then describes, using detailed examples, the building of actual databases used in reef fish stock assessment projects.

1. Define the data. Information should be stored in its smallest logical parts. Parameters defining all levels of sampling that are recorded on the data sheet must be recorded in a separate field in the database table. These include the identification for each individual replicate (smallest sampling unit) through to the level of sites, habitats, reefs and country. The variables that were actually measured at each replicate count (fish species, length, number, water depth, sea conditions, etc.) must also have their own field.

Defining the data also includes deciding on the appropriate data type. For example, it is infinitely easier to sort and group time information which has been recorded as date/time format rather than as text format, because properties of dates will not be recognised in the latter format.

Most database programs can accept long records, such as full species names, therefore the use of cryptic abbreviations should be avoided because they may cause confusion to others. The need to enter long names repeatedly into a data table can be avoided by creating a *reference table* as outlined in the next step.

2. Group fields into separate tables. The aim of this process is to reduce the amount of data that has to be repeated, resulting in a more efficient use of computer space as well as faster data entry and retrieval.

This process requires some consideration of the design of the field sampling. In the example of a UVC survey, each individual replicate count has only one set of information relating to its location, date, weather conditions, etc. In contrast there are often many records of fish within each replicate. All fields which relate to the replicate description (time, observer, depth, etc.) should therefore be grouped in one table so this information is recorded only once (i.e. only one record per replicate). The replicate data (fish species observed, lengths, etc.) should be recorded in a separate table, along with a field common to both which can be used to *link* the separated data together (see step 3 below). This avoids recording redundant sample identification information.

The use of separate, but linked, tables can also assist in entering long or complex information. For example, a reference table containing the full Latin names of all target species can be linked to the data table by suitable abbreviations. Thus, where there is repetitive recording of species names, as in the replicate data table, only the abbreviations need to be entered. A suitable link (step 3 below) to a species list reference table will enable the full name to be used in subsequent reports or by analysis

packages. An additional advantage of this system, compared with using abbreviations or codes alone, is that an unfamiliar (or forgetful) user can always check the full names in the reference table, which makes the whole database more self-explanatory and user-friendly.

Refining fields

For a relational database to work efficiently, each database table should have a field defined as the *primary key*, in which each record has a different value and is thus uniquely identified. This assists in the sorting of data. The primary key field cannot contain duplicate records. In some cases there will be an existing field which has unique records and this should be used as the primary key. Examples of such fields include sample identification numbers in a replicate description table and the species abbreviations field in a reference table of full species names. In a table with no unique records a primary key field should be added.

3. Define relationships between tables. Defining relationships between tables is essential when setting up a relational database. It is possible to initially set up the required tables with no defined links between them; however, this means that relationships must be re-defined every time data are extracted from more than one table during analysis. This is inefficient, can lead to errors due to inconsistencies in the links and does not fully utilise the relational capabilities of the database. As there is only one correct relationship structure between any set of tables in a database there is no advantage in leaving tables unlinked.

In order to define relationships tables must share a *linking field* - a field that is common to each table in the link. This allows the data from separate tables to be brought together in a logical way.

The most common linking relationship is one in which each record in one table is unique, but relates to numerous records in another table. For example, in a UVC database the records in a replicate description table (e.g. *Replicate Identification* table, see 5.3) are unique (there is only one record for each replicate), but there may be many records for each of the counts in the table containing the survey results (i.e. fish identification, numbers and lengths, *see Fish Count Data* table, section 5.3). In such cases the link must be defined as *one-to-many*. In a *one-to-one* link each

record from one table relates uniquely to only one record from another table. In most of these cases it is more efficient to combine this information into a single table. A *many-to-many* link is ambiguous and therefore cannot be established in a relational database.

Refining relationships

To avoid ambiguity in a *one-to-many* link it is essential that each record in the common linking field on the *one* side is unique. It is therefore a good idea to define the linking field in this table as the *primary key* so duplicate records are automatically avoided.

When establishing links between tables in a database it is advisable to *enforce referential integrity*. This means that any record entered into a table on the *many* side of a link in which the value in the linking field is not the same as that of records already present in the table on the *one* side of that link, is automatically rejected by the computer. In this way many mistakes, such as entering incorrect species abbreviations, can be eliminated from the database at an early stage.

Once the basic structure of the database has been established some data should be entered to test that all information handling requirements can be met. There are many more refinements which allow databases to work more efficiently, especially with respect to data entry, such as *forms*. These procedures are best found in the software manual of your particular database.

5.3 Creating raw data tables for UVC data

In this example the database contains three linked tables:

Fish Count Data: contains the data on the species, size and number of fish observed.

Replicate Identification: a reference table to store the information that identifies the characteristics of each replicate.

Species List: a reference table for the full Latin names and biomass calculation details of each species of fish.

5.3.1 Creating a new data table (based on Microsoft Access 2)

Double click the Microsoft Access icon.

A relatively empty screen is presented.

Click the mouse cursor on the word File.

A file menu appears with a list of options which allow you to create, open, repair and save database files.

Click New Database.

A New Database dialogue box appears where you enter a file name.

Enter *UVC* (as an example). Click OK.

Your screen should look like the above, showing the Database box.

5.3.2 Creating the Replicate Identification table

Click the New button in the database box.
Click New Table in the New Table box.

The Table design box will appear where you will need to define the field name, its data type and description.

Enter *Sample ID* then press the ← key.
Click the mouse on the ← box and then click *Number*.

In the *Sample ID* field each different sampling unit (replicate) is identified by a unique number. The data type *Text* appears - you need to change the type to *Number*. You have now set a field name and its corresponding data type. You should enter a description in the next column, e.g. identifies replicate.

Continue to enter field names and data types for the remainder of the table and add a description for each field.

Examples are given in the table below. Additional fields containing more information about each replicate (e.g. water depth, observer, etc.) may be added to this table.

The *primary key* needs to be set on the *Sample ID* field. This field uniquely identifies each replicate and is the linking field between tables. Click the mouse in the *Sample ID* field. Click the Edit menu on the top of your screen. Click Set Primary Key.

The *Sample ID* field will have a key symbol next to it and the table design box should now look like that shown below.

To reduce errors in data entry it is recommended that each field is defined in terms of size, validation rule, format and decimal places. The settings are entered in the field properties box beneath the Table design box.

You have now completed the design of the *Replicate Identification* table. You now need to name and save the new table.

Click the File menu. Click Save As. Enter: *Replicate Identification* in the Table Name box and click OK.

The design of the table is complete and data entry can now proceed.

Click the Datasheet View icon and enter the data (as seen in Table 5.1).

If you wish to alter the design of your table or query at any time use the *Design* icon to re-enter the design screen.

Databases

5.3.3 Creating the Fish Count Data table

Repeat the above steps to open a new table. Enter field names and data types for the table as set out below.

This table structure is efficient in terms of space saving and data entry because it groups individuals of the same species and lengths into a single record: the total number of individuals in each size group is entered into the *Frequency* field. In this example an *ID* field has been added for the purpose of setting a primary key in this table with a *counter* data type which will automatically enter a value to identify each record uniquely.

Name and save the new table as *Fish Count Data*.

5.3.4 Creating the Species List table for species names and length-weight relationships

The *Species List* table acts as a reference list for the full species names and also stores the information needed to calculate biomass. The species abbreviations used here are suggested standards only, the main requirement being that they are unique. Weights of fish are calculated from length estimates derived from UVC surveys by using species-specific length-weight relationships (Chapter 3). A table is required listing all *length-weight constants a* and *b* for each species of fish in the census list (Table 5.1). The constants were obtained primarily from Kulbicki et al (1993), and also Wright and Richards (1985). UVC fish length estimates are in cm. Length-weight relationships may be published in cm (Kulbicki et al 1993) or mm (Wright and Richards 1985). Thus, the UVC estimates must be standardised or made compatible with the length-weight relationship. For the Wright and Richards (1985) constants, the UVC length estimates need to be multiplied by 10. This is achieved by the extra field called *length conv*. If it is desirable to summarise data at a level other than the taxonomic groups listed here (e.g. on the basis of trophic groups) a new field containing this information should be added to this table.

Repeat steps in 5.3.1 to open a new table. Enter field names and data types as set out below.

Field Name	Data Type	Description
Family	Text	field used for sorting or grouping by Family
Species	Text	field used for sorting or grouping by Species
Kul/WR Species	Text	Publication Source: Kul=Kulbicki et al 1993; WR=Wright & Richards 1985
length conv	Number	WR equation uses mm therefore UVC length (cm) must be converted by x10
a	Number	length weight constant
b	Number	length weight constant
Species Abbreviation	Text	unique species name abbreviation used in data tables

Note: **Kul/WR species** field provides the publication source for each species (no 22 etc. = reference source in Kulbicki et al 1993).

The primary key for this table should be set on the **Species Abbreviation** field and therefore the requirement for unique records will be automatically enforced.

Click the mouse cursor on the *length conv* field and click the Validation Rule box and enter 1 or 10. Name and save the new table as Species List.

To avoid data entry errors in the **length conv** field a data validation rule is specified. A validation rule for the **length conv** field is set so that only 1 or 10 can be entered.

5.4 Creating data tables for creel and questionnaire surveys

The basic structure of the databases for the creel and questionnaire surveys are similar to the UVC database.

The creel survey database consists of three raw data tables:

(i) **Creel Survey Catch** stores the data on the catch, such as species, numbers and weights. The analysis of data from this table will provide information on **Catch** (see Chapters 4 and 6).

(ii) **Creel Survey Respondents** stores all the sampling (replicate) data, such as date, time, area fished, creel survey number, boat, gear etc. and is linked to the **Creel Survey Catch** table by the **Sample ID** field. The analysis of data from **Creel Survey Respondents** will provide information on **Fishing Effort** (see Chapters 4 and 6).

(iii) **Species List** provides full names of fish species recorded and is linked to **Creel Survey Catch** table via the **Species Abbreviation** field. This table may be the same as that used in the UVC database (see 5.3.4 above - the length-weight conversion information can be ignored), although additional species may have to be added to account for all fish observed by this different survey method.

Table 5.1 The *Species List* table which lists those species used in the ACIAR/DPI UVC Project (Samoilys et al 1995). Constants **a** and **b** refer to the length-weight relationship Weight = a·Length^b. (no 22 etc. = reference source in Kulbicki et al 1993).

Family	Species	Abbreviation	Kul/WR species	length	a	b
Acanthurid	*A. D+M+X*	A dmx	Kul-*A. dussumieri*	1	0.011	2.761
Acanthurid	*A. lineatus*	A lin	Kul-*A. lineatus 22*	1	0.0192	3.072
Acanthurid	*A. nigricauda*	A nig	Kul-*A. nigricauda 7*	1	0.08	2.61
Acanthurid	*A. triostegus*	A tri	Kul-*A. triostegus*	1	0.052	2.394
Acanthurid	*Acant+Ctenot+Zebras.*	A sp	Kul-*Zebrasoma veliferum*	1	0.0471	2.857
Acanthurid	*C. striatus*	C str	Kul-*C. striatus*	1	0.0278	2.997
Acanthurid	*N. brevirostris*	N bre	Kul-*N. brevirostris*	1	0.0102	3.128
Acanthurid	*N. hexacanthus*	N hex	Kul-*N. brevirostris*	1	0.0102	3.128
Acanthurid	*N. tuberosus*	N tub	Kul-*N. unicornis*	1	0.0222	2.988
Acanthurid	*N. unicornis*	N uni	Kul-*N. unicornis*	1	0.0222	2.988
Acanthurid	*Naso spp.*	N spp	Kul-*N. brevirostris*	1	0.0102	3.128
Labrid	*Ch. fasciatus*	C fas	Kul-*Cheilinus chlorourus*	1	0.062	2.778
Labrid	*Ch. trilobatus*	C tri	Kul-*Cheilinus chlorourus*	1	0.062	2.778
Labrid	*Ch. undulatus*	C und	WR-*Bolbometapon muricatum*	10	0.0000082	3.1
Labrid	*Choerodon spp.*	C spp	Kul-*Choerodon graphicus*	1	0.00895	3.153
Labrid	*H. fasci + melas*	H f+m	Kul-*Cheilinus chlorourus*	1	0.062	2.778
Lethrinid	*Gymnocranius spp.*	G spp	Kul-*Gymnocranius japonicus*	1	0.0288	2.959
Lethrinid	*L. harak*	L har	Kul-*L. harak*	1	0.0154	3.043
Lethrinid	*L. nebulosus*	L neb	Kul-*L. nebulosus*	1	0.0265	2.943
Lethrinid	*L. olivaceus*	L oli	Kul-*L. olivaceus*	1	0.0662	2.78
Lethrinid	*L. xanthochilus*	L xan	Kul-*L. xanthochilus*	1	0.0378	2.872
Lethrinid	*Lethrinus spp.*	Leth	Kul-*L. nebulosus*	1	0.0265	2.943
Lethrinid	*M. grandoculis*	M gra	Kul-*M. grandoculis*	1	0.0259	2.989
Lutjanid	*Aprion virescens*	A vir	Kul-*Aprion virescens*	1	0.0351	2.869
Lutjanid	*Lutjanus spp.*	Lutj	Kul-*L. argentimaculatus*	1	0.064	2.761
Lutjanid	*L. bohar*	L boh	Kul-*L. bohar*	1	0.0175	3.019
Lutjanid	*L. carponotatus*	L car	Kul-*L. fulviflammus*	1	0.0257	2.936
Lutjanid	*L. fulvi+ehren*	L f+e	Kul-*L. fulviflammus*	1	0.0257	2.936
Lutjanid	*L. fulvus*	L ful	Kul-*L. fulvus*	1	0.0275	2.937
Lutjanid	*L. gibbus*	L gib	Kul-*L. gibbus*	1	0.021	2.996
Lutjanid	*L. kasmi+quinq*	L k+q	Kul-*L. quinquelineatus*	1	0.0244	2.959
Lutjanid	*L. monostigma*	L mon	Kul-*L. russelli*	1	0.0327	2.85
Lutjanid	*L. rivulatus*	L riv	WR-*L. rivulatus*	10	0.000015	3.05
Lutjanid	*L. russelli*	L rus	Kul-*L. russelli*	1	0.0327	2.85
Lutjanid	*L. semicinctus*	L sem	Kul-*L. fulviflammus*	1	0.0257	2.936
Lutjanid	*Macolor spp.*	M spp	Kul-*L. bohar*	1	0.0175	3.019
Scarid	*B. muricatum*	B mur	WR-*Bolbometapon muricatum*	10	0.0000082	3.1
Scarid	*Cetoscarus bicolor*	C bic	WR-*Scarus harid*	10	0.000018	2.99
Scarid	*Hipposcarus longiceps*	H lon	WR-*Scarus harid*	10	0.000018	2.99
Scarid	*S. altipinnis*	S alt	Kul-*S. altipinnis 22*	1	0.0233	2.98
Scarid	*S. frenatus*	S fre	Kul-*S. altipinnis 22*	1	0.0233	2.98
Scarid	*S. ghobban*	S gho	Kul-*S. ghobban*	1	0.0141	3.061
Scarid	*S. microrhinos*	S mic	Kul-*S. gibbus 22*	1	0.0388	2.897
Scarid	*S. niger*	S nig	Kul-*S. altipinnis 22*	1	0.0233	2.98
Scarid	*S. rubroviolaceus*	S rub	Kul-*S. rubroviolaceus 22*	1	0.0136	3.109
Scarid	*Scarus spp.*	S spp	Kul-*S. sordidus*	1	0.0319	2.927
Serranid	*Anyperodon leuco*	A leu	Kul-*Epinephelus areolatus*	1	0.0154	2.977
Serranid	*C. argus*	C arg	Kul-*C. argus*	1	0.0155	3.022
Serranid	*C. cyanostigma*	C cya	Kul-*C. boenak*	1	0.0106	3.081
Serranid	*C. miniata*	C min	Kul-*C. miniata*	1	0.0655	2.757
Serranid	*E. caeruleopunctatus*	E cae	Kul-*E. caeruleopunctatus*	1	0.0257	2.913
Serranid	*E. maculatus*	E mac	Kul-*E. maculatus*	1	0.0255	2.899
Serranid	*E. polyphekadion*	E pol	Kul-*E. microdon*	1	0.0257	2.923
Serranid	*P. aerolatus*	P aer	Kul-*Plectropomus leopardus*	1	0.00923	3.078
Serranid	*P. laevis*	P lae	Kul-*Plectropomus leopardus*	1	0.00923	3.078
Serranid	*P. leopardus*	P leo	Kul-*Plectropomus leopardus*	1	0.00923	3.078
Serranid	*P. maculatus*	P mac	Kul-*Plectropomus leopardus*	1	0.00923	3.078
Serranid	*P. oligacanthus*	P oli	Kul-*Plectropomus leopardus*	1	0.00923	3.078
Serranid	*Variola spp.*	V spp	Kul-*Variola louti*	1	0.0134	3.036

Databases

5.4.1 Creating the Creel Survey Respondents table for the creel database

Repeat steps in 5.3.1 to open a new table. Enter field names and data types as shown in the example below.

The *Sample ID* field should be set as the primary key. Additional information recorded in the ACIAR/DPI UVC Project is not shown in the table above because it was not used in analysis (Samoilys et al 1995). This information could have been stored in this table within additional fields relating to: Time; Boat (Solomons) or Licence/boat (Fiji); Recorder; Did you catch these fish? (Fiji only); Fisher; Have you captured any of these fish in the study area? (Fiji only); Landing (place), etc.

5.4.2 Creating the Creel Survey Catch table for the creel database

Repeat steps in 5.3.1 to open a new table. Enter field names and data types as set out below.

Field Name	Data Type	Description
Sample ID	Number	Link to respondents table
Species Abbreviation	Text	Link to Species List table
Fish Number	Number	Number of fish in catch
Weight	Number	Weight of fish in catch
D	Counter	unique records for primary key

As in the *Fish Count Data* table in the UVC database, if a primary key is desired in this table the *ID* field must be added because records in existing fields will not be unique. Codings for fishing gear used in the questionnaire and creel databases can be linked to a further reference, or look-up table, so the full names of fishing methods can be used in data manipulation and analysis.

5.4.3 Creating the Last Trip and the Last Trip Catch tables for the questionnaire database

The examples given here refer to the last section of the questionnaire data sheets: section 4: Catch (see Chapter 4), which collects data on the fisher's most recent fishing trip. The questionnaire survey database consists of three raw data tables:

(i) Last Trip Catch: this table provides the data on the catch itself, such as species, numbers and weights.

(ii) Last Trip: this table provides all the sampling data, such as date, time, area fished, questionnaire survey number, boat, gear etc., and is linked to the *Last Trip Catch* table by the *Sample ID* field.

(iii) Species List: provides full names of fish species recorded, and is linked to *Last Trip Catch* table via the *Species Abbreviation* field. Again, this table may be the same as that used in the UVC and creel databases, providing all species observed have been included.

Repeat steps in 5.3.1 to open a new table. Enter field names and data types for the *Last Trip* table as shown in the example below. The primary key is set on the *Sample ID* field.

Field Name	Data Type	Description
Sample ID	Number	link to catch data table
Survey Number	Number	1, 2 or 3
Date	Date/Time	date of fishing trip
Area fished	Text	area fished
Crew	Number	Number of Crew
Gear	Text	type of fishing gear
Trip Length	Number	fishing trip length in hours

Enter field names and data types for the *Last Trip Catch* table as shown below.

Field Name	Data Type	Description
Sample ID	Number	link to last trip data table
ID	Counter	primary key record identifier
Species abbreviation	Text	link to species list table
Fish numbers	Number	number of fish caught
Fish weight	Number	weight of fish caught

5.5 Linking tables

The relationships between the tables in a database need to be defined. The following steps describe this process for the UVC database.

Click on the Edit Menu and select the Relationships option. Click on the Relationships Menu and select the Add Table option. From the pop-up Add Table box which appears, select each table in the database in turn and click the Add button to place them in the empty relationships design box. Close the Add Table menu.

The relationship between the data tables needs to be established as *one-to-many*, because there are many records in the *Fish Count Data* table, (on the *many* side), linking

to each record in the ***Replicate Identification*** table (the ***one*** side).

> **Drag the Sample ID field in the *Replicate Identification* table box to the Sample ID field in the *Fish Count Data* table box.**

A ***Relationships*** box appears in which the join properties need to be defined.

> **Select *Enforce Referential Integrity* (see section 5.2.1) and then the *one-to-many* option. Click the *Join Type* button.**

The link can be defined further according to the hierarchy of the tables. In the UVC database structure all data in the ***Fish Count Data*** table pertain to records from the ***Replicate Identification*** table. To reflect the survey design accurately all information from replicate records should be displayed along with those count data which correspond to these replicates. With data linked this way there will be missing values for replicates where no fish were observed. Because these missing values are nulls, not zeros, they will be ignored in any data calculations based on the ***Fish Count Data*** table. This structure will serve as a reminder that ***all replicates must be taken into account in the calculation of averages***, etc.

> **Select the join option which includes all records from the *Replicate Identification* table and only those records from the *Fish Count Data* table where the joined fields are equal. Click on the Create button; the link is now established.**

A line now connects the tables with symbols for ***one*** (1) and ***many*** (⇐) at the appropriate ends and an arrow pointing towards the ***Fish Count Data*** table.

> **Link the *Species List* table to the *Fish Count Data* table by the Species Abbreviation field using a similar *one-to-many* join procedure.**
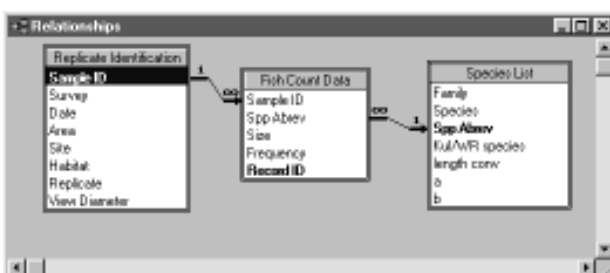
The join type should include only those records from ***Species List*** where a corresponding species is recorded in ***Fish Count Data***. The relationships between the tables should appear similar to that shown above.

Join types can be modified at any time by double clicking on the connecting line.

Exactly the same procedures are used to link the tables in the creel and questionnaire databases.

5.6 Database management

The ongoing operation and management of a database is as important as its initial set-up. Operation and management involves data checking procedures, back-up procedures and established protocols for data handling. The chapter by Bainbridge and Baker (1994) in the AIMS manual provides a thorough description of these procedures, which should be documented in a ***SOP (standard operational procedure).*** Here we cover some data checking procedures.

Clearly, a database is only good if the information from the data sheets has been entered correctly. This rather obvious statement is made because the mere existence of data in a database can give it a false sense of validity. This is especially true if various operators are performing different tasks; for example, data entry is done by one researcher, but data manipulation and analysis is done by another. The management of a database must include procedures, outlined in the SOP, that ensure data are entered correctly. One method of data validation available in Access is to use customised forms for data entry. Forms impose conditions on the type of data which can be entered into each field and can have built-in prompts to help the user. However errors in data entry which are not detected by the data validation conditions set in Access can still occur, mainly through human error. Database errors can be corrected by re-entering data or by using the update query function in Access (see your database software user's manual for more details on the use of forms and update queries).

5.6.1 Data checking

The need to check data immediately after it has been entered into the database is an important part of maintaining data quality. Standard procedures for checking data are as follows:

Databases

1. Print the data and then check the print-out against the data sheets. Data checking requires two people to save time and decrease the likelihood of errors. A simple directive would be that all data is entered by one researcher and must be checked at least once by a second person.

2. Mark errors on the print-out and then update the database with the corrections.

3. File corrected print-outs of the database records and the raw data sheets, backup the database and store it as an archive. Preferably one archive copy should be stored off-site.

The dynamic dataset (dynaset) presented by one query can in turn be examined, remanipulated and sorted by another query. This can be a useful feature in summarising and describing data at various levels of sampling, i.e. the mean number of each species can be re-examined at the level of family or trophic group by making a new query using the data from the species summary query.

This section describes in detail the basics of creating a simple query involving linked tables and the calculation of new information from the available data. More sophisticated queries which are used to present statistical summaries of data at different levels of sampling are also discussed, with examples from actual fisheries databases.

5.7 Data manipulation

There are many approaches to the manipulation of data. Because a database makes it so easy to extract information, any non-systematic approach to data handling will inevitably lead to a plethora of new "slightly improved" data sets of uncertain vintage, a situation which leads to confusion. It is therefore of paramount importance to design a systematic approach to data handling before any information is processed.

A useful way to think about the examination of data is to use the tables as the "unalterable truth" from which all subsequent information must be derived. On the whole it is best to avoid creating new tables in a database unless it is for the purpose of adding genuinely new information.

Creating a database *select query* is probably the most useful method of extracting and summarising data. A query does not write the selected data permanently to a table but presents it arrayed as a *dynaset*, a virtual table which is recalculated each time the query is opened or run. This means that as new data is entered into a data table, or errors are corrected, these are automatically reflected in the query results.

In a relational database, queries usually need to examine more than one table simultaneously in order to extract sufficient information to perform the required summaries and calculations from the data. If the database has been set up properly, links between the various data tables will be automatically transferred to any query where those tables are examined. It is possible to add and customise links between any tables in a query.

5.7.1 Creating a query from data tables

In the UVC database created previously, fish weights must be calculated from lengths recorded in the *Fish Count Data* table using information stored in the *Species List* table. This is achieved by creating a query which examines both of these tables, performs the relevant length-based weight calculation and places the biomass figure in a new field.

Creating the query for length-weight conversion

Click the Query button in the Database box. Click the New button. Click New Query in the New Query box.

An Access query has been created and an Add Table box appears from which you can select the appropriate tables.

Double-click the *Species List* table. Double-click the *Fish Count Data* table. Click the Close button.

Joins, reflecting those already made when table relationships were defined (section 5.5), should automatically appear between the tables. Now the fields need to be selected for the query.

In the *Fish Count Data* table box double-click on the fields: Sample ID, Size and Frequency. In the *Species List* table box double-click on the fields: Family and Species.

Each column in the lower window will contain a selected field and the *Show* row will contain a crossed box, indicating that the field will be displayed. A new field needs to be created in which fish weights will be calculated.

Move to the first blank column in the query and click in the field row (this will be next to the *Frequency* field) and enter: Weight: ([Frequency]*[a]*(([Size]*[length conv])^[b]))

This equation calculates biomass from length (size) using length-weight relationships (see Chapter 3) and names the field *Weight*. The query is now complete and should look like the screen view below.
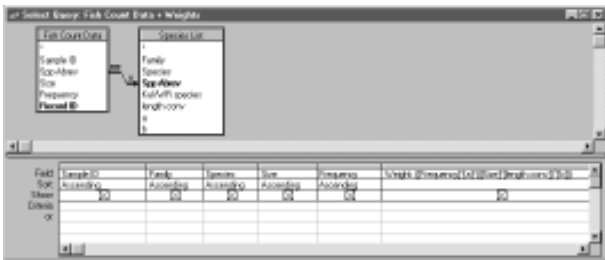
You should save the query.

Click the File menu. Click Save As. Enter: *Fish Count Data + Weights.*

It is important to remember that the weight calculation differs for each species and individual size. The separation of the data by Sample IDs retains the biomass data summarised at the level of replicate for future comparison between survey areas, sites, etc. Subsequent summaries must therefore use the results of this query as their starting point.

5.7.2 Data summaries - density and biomass calculations

When summarising data standard units must be used. For example a typical visual census may only be 250m^2 in area (e.g. 50m x 5m transects). Alternatively the areas of point counts will vary if the observer measures the radius after the count (Chapter 3). Numbers and weights of fish in the UVC data must be summarised per standard area of reef to give densities and biomass for each replicate or census (see Chapter 3). The area standard is usually 1000m^2. It is only after this has been achieved that calculations of means and variances from any level of sampling can be calculated.

In the manipulation of data you must know what end result is required, and this process depends in part on the questions being asked (see Chapters 2 and 6). For example are summaries required at a family or species level, or even a trophic level (e.g. predators, herbivores, planktivores)?

Usually the aim of any data manipulation for statistical work is either to summarise the data (at the desired level) as a series of replicate values (e.g. for use in statistical procedures such as ANOVA, see Chapter 6) or to obtain an average of values across a sample of these replicates. Therefore, it is usually the first step in any analysis to summarise the data at the lowest level of replication of the sampling design. With an Access database this usually involves the creation of an initial query which groups data at the level of each replicate sampling unit. It should then be possible to use this initial query to extract and process data in a consistent way for any combination of survey sites, regions, or individual replicates.

From this point it would seem to be a simple process to quickly calculate averages, standard deviations, etc. using statistical functions which are built into most database programs. However, because of the way information is recorded in a relational database careful thought is required in the manipulation of data for statistical procedures. This important point arises because only those fish actually observed are recorded in the database, even though in some instances during a UVC or creel survey there may have been a great deal of sampling effort expended where no fish (of a particular category) were observed. This has important implications for a scientific survey. It is obvious that a count of zero fish in a sample has as much relevance as the observation of any other number of fish. Operations involving statistics which are based only on those values recorded in the database would ignore these zero counts, therefore returning erroneous results. This means that there can be serious limitations to the usefulness of the built-in summary statistics functions in database programs.

To avoid this problem by entering all the zero counts for every target species in every replicate count would be time (and computer space) consuming. Fortunately there are versions of the formulae for means and variances which use only the sums of replicate values, therefore zero values are not required. If desired statistical quantities (mean, variance, etc.) are calculated by specifying these formulae in customised queries, the missing zero counts have no effect on the outcome. It is necessary, however, to know the number of samples from which these totals were derived. In some sampling methodologies, such as UVC, where the number of replicates is always constant, this known value can be inserted directly into the specified statistical formulae. In most other sampling such as CPUE surveys

(creel and questionnaire), the replication may vary between areas, so a single figure cannot be entered into the formulae.

In summary, not all replicate sampling units will appear in a summary of data except in rare cases where, for each category of fish in each survey area examined, there are observations in every replicate. Where there is a variable number of replicates the actual number conducted must be obtained via an extra procedure that counts the sampling units in the *Replicate Identification* table. This highlights why it is so important to record details of every sampling unit in this table.

5.7.3 UVC data

Creating an initial query to calculate fish density and biomass per replicate

This step in data manipulation summarises fish numbers and the newly calculated biomass data at the level of each replicate. The following example is customised to give the total number of fish per replicate and the total weight of fish per replicate, grouped at the family level. Grouping could be based on the species or trophic group level by substituting the appropriate field for Family.

Create a new query. Add table: *Replicate Identification*, Add query: *Fish Count Data + Weights*

The join type must be customised as the relationship between the query *Fish Count Data + Weights* and the other tables has not been previously specified. Sample ID is the common linking field. Records of all the sampling effort information from *Replicate Identification* should be presented along with only those corresponding records from the query *Fish Count Data + Weights* where the joined fields are equal.

Drag the following fields into the new query from the *Replicate Identification* table: Survey, Habitat, Site, Sample ID, View Diameter. Drag the following fields into the new query from the *Fish Count Data + Weights* query: Family, Frequency, Weight. Sort the fields: Survey, Habitat, Site, Sample ID and Family in ascending order by moving to the Sort row and Click the ⊷ box. Click Ascending.

Future tables should be presented in an ascending order to provide consistency and uniformity.

Click the Group By icon.

All fields should have a *Group By* status in the *Total* row. This arranges the data within the grouped fields so that each combination (of Sample ID within Site, Habitat and Survey) will only appear once. The numbers and weights of fish in these specified groups need to be summed.

Move to the Frequency field's Total row. Click the ⊷ box. Click Sum. Repeat summing commands for the Weight field.

A new field to calculate the area of each replicate circular point count needs to be created, using the formula:
$Area = \pi r^2$.

Move to the first blank column in the query and click the Field row. This will be next to the weight field. Enter Area: 3.14159*(([View diameter]/2)^2)

In the Total row Click the ⊷ box and select *Expression*.

New fields such as Area, which use existing data to calculate values, are *Expressions*, and need to be defined as such. The query so far should resemble that shown below.
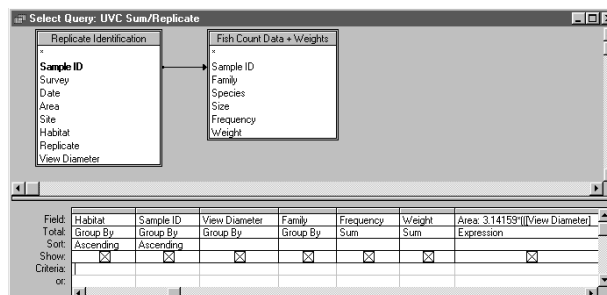
Create new fields for density and biomass to calculate numbers and weights of fish per 1000m² area. The formulae used to create density and biomass fields are as follows:
Density: 1000*([SumOfFrequency]/[Area])
Biomass: ([SumOfWeight]/[Area])

Note: the biomass formula also converts weights from grams to kilograms.

For the subsequent calculation of variance, the density squared and biomass squared (x^2 in statistical formulae) for each replicate must also be calculated and stored in additional fields. The formulae used for these fields are as follows:
Sq Den: [Density] ^2
Sq Biom: [Biomass]^2
Save Query as *UVC Sum/Replicate*.

This is the initial summary of density and biomass at the level of individual replicate. If greater specificity is required, for example, if we wish to select only records of the family Lutjanidae from the database, enter in the *Criteria* row: "Lutjanid".

The results of this query are the starting point for subsequent grouping or condensing of data. Some statistical procedures, such as ANOVA, require the data in this format for analysis. The results of the query may then be copied directly to the desired program for analysis (e.g. Excel, see Section 5.9). It must be remembered however, that at this stage of data manipulation only those replicates in which fish were observed will appear in the query results. Creating a procedure to correct for this would be complex, and given that a large number of species are usually grouped together in the analysis, it is usually quicker to add the missing zero values manually when the data have been copied into a spreadsheet.

Creating the UVC Mean & Variance query

Means and variances for both density and biomass can be calculated in new expression fields using data from existing fields in the *UVC Sum Replicate* query.

Create a new query. Group By fields: Family, Survey, Habitat, Area and Site; Sum on fields: Density, Biomass, Sq Density and Sq Biomass.

The fields for calculating means and variances need to be created. Mathematically, the formulae for the expressions are as follows (see also Chapters 3 and 6):

Mean, $\overline{x} = \dfrac{\sum x}{n}$

Variance, $s^2 = \dfrac{\sum x^2 - \dfrac{(\sum x)^2}{n}}{n-1}$

In Access the mean and variance are written (taking 12 to be n, the standard number of replicates in UVC) as follows:

Mean Density: [SumOfDensity]/12

Var Density: ([SumOfSq Den]-(([SumOfDensity]^2)/12))/ 11

Mean Biomass: [SumOfBiomass]/12

Var Biomass: ([SumOfSq Biom]-(([SumOfBiomass]^2)/12))/ 11

5.7.4 CPUE data

The analysis of catch data is usually more complex than that of UVC data, as logistical and practical constraints of field work often mean that there are not an equal number of replicates in each sampling area. The total fishing effort represented in the creel or questionnaire surveys must be derived from the data table that stores information from the survey respondents. Procedures broadly similar to those detailed for UVC data can be used to summarise the catch totals per replicate and these can be combined with effort information to obtain values of CPUE. The following example is based on creel survey data, and summarises catch, effort and CPUE per species. Questionnaire data can be treated in the same way.

Calculation of CPUE for each replicate

An initial query summarising numbers and weights per replicate must be created.

Construct a query based on the *Creel Survey Respondents*, *Creel Survey Catch* and *Species List* tables. These tables should be automatically linked together by appropriate fields. Use the Group-By function (see section 5.7.3) to summarise the data by Survey, Area, Site, Gear, Sample ID, Effort, Crew and Family; use the Sum function to calculate totals in the Numbers and Weight fields. Catch and effort information is calculated in two new expression fields:

CPUE (Number): [SumOfFish Frequency]/ ([Effort]*[Crew])

CPUE (Weight): [SumOfWeight]/([Effort]*[Crew])

The query should look similar to that shown below.

Databases
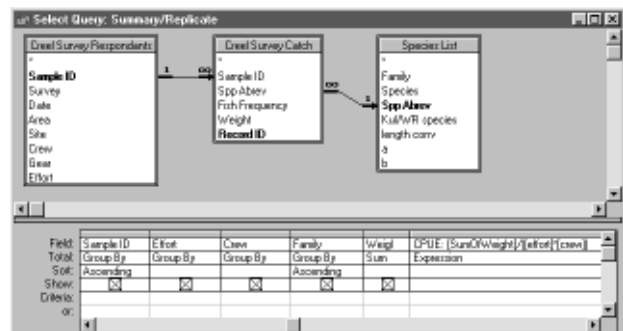
Calculation of replication in surveys

A query is needed to calculate the number of sampling units undertaken at whatever level the survey is be analysed. This query will be combined with the results of the previous query, therefore, grouping must be based on the same fields as were summarised previously. In the example provided the replication of the catch survey has been summarised per Survey, Area, Site and Gear from the *Creel Survey Respondents* table.

Create new query. Add and Group By fields: Survey, Area, Site and Gear; Add and Count on the Sample ID field.

The query will return a count of the number of replicates for each gear in each group of sites within areas within surveys. The finished design should look similar to that shown below.
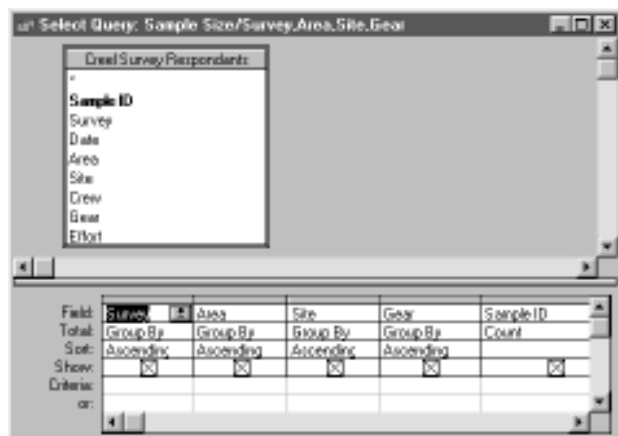
Calculating mean CPUE from catch data

The information on the amount of replication from the *Sample Size/Survey Area, Site, Gear* query above, together with the calculations of CPUE for individual replicates (*Summary/Replicate* query), can be used to derive mean CPUE for a specified level of grouping.

Create a new query based on the *Summary/Replicate* query and the *Sample Size/Survey,Area,Site,Gear* query. Link queries by all fields which were used to group the data. Create new expression fields to calculate means and variances (see UVC queries in section 5.7.3).

It is essential that linkages are specified on those fields which were used to group data in both tables. This means that any changes to the level of grouping in one query must be accompanied by corresponding changes in grouping in

the other query. Consequently, the links will need to be modified. In the example provided the final query design for calculation of mean and variance should look like that below.
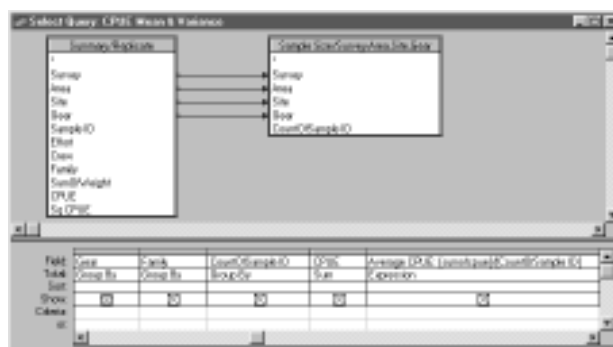
This section has provided some introductory examples of data processing using queries in Access. The applications are very flexible and can execute much more sophisticated queries for data analysis than have been covered here. For more information refer to the database software user's manual.

5.8 Summarising data by crosstab queries

Storing information in a database format as used in the preceding examples can make it difficult to gauge trends and make comparisons among the data. This can be overcome by using one of Access's features which quickly presents a large amount of summary data in an easily readable spreadsheet or table format. These summaries are known as *crosstab queries* and can be used as the basis of advanced data analysis or reports. For example, a crosstab query can be designed to show the total numbers or average density for each species per site or area.

5.8.1 Producing a crosstab query

A *crosstab query* can be created either by using the Query Wizard and following the directions in the dialogue boxes, or by custom design. The process for designing a crosstab query is similar to designing a select query (section 5.7), except you must specify which field(s) are to be used for *row headings*, *column headings* and *the value*.

There are two important points to note about using a crosstab query:

(i) Data returned is a snapshot, a type of recordset that is not updatable. For this reason crosstab queries should only

be used for final summaries, analysis or reports after all data has been entered and checked.

(ii) The summary statistics functions in the crosstab query *value* field assume that all replicates are represented in the data source. Thus, if there are zero values for some replicates they must be included in the dataset. This is important to remember, particularly for UVC data, because there are often zero values for certain species or families that are not entered in the original raw data tables in Access. It is therefore best not to use *crosstab queries* to calculate summary statistics, but only to use them to present existing data in a more readable format. In the example below the crosstab query is simply rearranging the summarised data in the *UVC Mean & Variance* query, and is *not* performing the summary statistics function specified in the Total row of the value field.

The design of a crosstab query which will summarise mean UVC-derived densities of all fish families across surveys, sites and habitats is shown below.
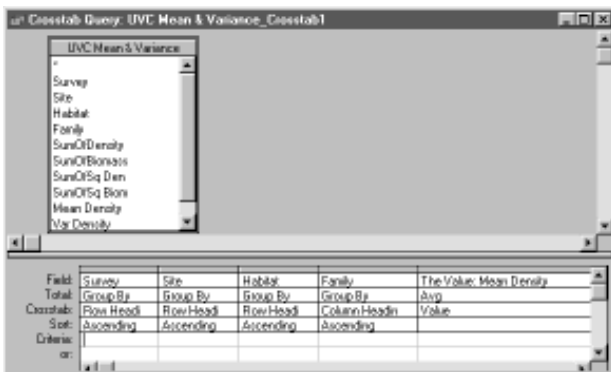
The crosstab query will display a table similar to that shown below.

Survey	Site	Habitat	Acanthurid	Labrid	Lethrinid	Lutjanid	Scarid	Serranid
A	1	Lagoon	2.64	16.76	7.45	1.06	48.88	7.86
A	1	Reef Slope	3.70	33.07	0.26	1.82	99.19	6.51

5.9 Importing data into Excel

Any table or query from Access can be imported into Excel provided that it does not exceed Excel's maximum number of records.

Click the grey button situated at the top left-hand corner of the first record of your query or table.

The query or table will be highlighted. If you wish to import a portion of your data drag the mouse cursor over the section of data desired and highlight it.

Click the mouse on the word Edit. Click Copy. Open Excel. Click the grey button situated at the top left-hand corner of an empty spreadsheet. Click the mouse on the word Edit. Click Paste.

The data from Access has now been transported to Excel as a spreadsheet and can be used as described in Chapter 6.

Databases

CHAPTER SIX: DATA ANALYSES AND APPLICATIONS

Melita Samoilys, Marcus Lincoln Smith and Iliavi Tuwai

6.1 Introduction

To analyse and interpret data obtained from UVC and CPUE surveys of coral reef fish stocks in order to assess the status of the fishery, a sequence of three basic procedures is recommended. This chapter provides guidelines and suggestions for each of these procedures.

(i) First, the data must be summarised to a standard form and examined graphically.

This first step provides the researcher with their first look at the data. The data should be summarised in a standard format (see Chapter 5, section 5.7.2) for comparisons across datasets and with similar data from other studies which use the same methodologies. It is highly recommended that these summarised data are also presented graphically to help visualise the data - it is useful at this stage to compare the graphs and familiarise oneself with the results while searching for any apparent trends in the data.

(ii) The second step addresses the questions posed in the original design of the research program.

This step will involve hypothesis testing and statistical tests which were identified when designing the sampling program (Chapter 2).

(iii) The third step involves stock assessment.

As discussed earlier (Chapter 1) there are a variety of assessment procedures for assessing fish stocks, but currently there is no clear consensus on which are the most suitable models for tropical reef fisheries. This manual does not cover the procedures for using surplus production models or yield per recruit models; readers should refer to the FAO manual (Sparre and Venema 1992) and to recent literature (e.g. Appeldoorn 1996) for further information on the models.

Yields, usually expressed as tons per square kilometre per year (t km^{-2} yr^{-2}) are the most widely used measure of a reef fishery and provide a useful basis for comparisons with other areas or fisheries. It is important to note, however, that the upper limit for sustainable harvests from reefs has not been accurately determined, and is likely to vary between areas (Dalzell 1996). This chapter describes simple calculations for estimating yields from UVC data to illustrate the potential information that may be obtained from fishery-independent surveys such as UVC surveys.

6.2 Descriptive data summaries

6.2.1 Summarising procedures

This section describes procedures for presenting means and standard errors of (a) stock abundance (density and biomass) of selected species groups from UVC data, and (b) catch, effort and CPUE from fishery creel and questionnaire surveys. It is very important that the mean (together with a measure of variability or dispersion around the mean), is used to present all estimates from the data. The mean is an estimate of the population, based on sampling the population (see Chapter 2). The spread of data around the mean, i.e. the dispersion or variability, is represented by the ***variance***, or ***standard deviation***, or ***95% confidence limits***. Refer to standard statistical text books such as Sokal and Rohlf (1981) and Zar (1984) for definitions of these terms; formulae are given below. Thus the mean and associated measure of dispersion indicate how confident we can be in the estimate. If the data are highly variable it will be difficult to detect differences in estimates of stock abundance, say, over time or between areas. A ***replicate*** is one of a set of measurements taken from the same area at about the same time. The mean is derived from a sample of replicates. Totals or cumulative totals are sometimes reported, such as the total number of fish counted at a site, or on a survey. On their own, they are largely meaningless values since they sum all the replicates and provide no measure of dispersion; they are more a measure of one's sampling effort than a measure of stock abundance! However, with Catch and Effort surveys, totals are often calculated to provide estimates of Total Catch or Total Effort for a certain area, or year, assuming the data collected provide a reasonable measure of such totals (see Chapter 4).

To summarise fish survey data, first decide on the levels within the survey at which the data are to be summarised. For example, there might be two levels: fishing pressure (or different areas), and time of survey. The sample mean

is then calculated by summing the replicates within a level and dividing by n, the total number of replicates (or sample size).

$$\text{Mean, } \bar{x} = \frac{\sum x}{n}$$

The variance about the mean is then calculated from the following formula:-

$$\text{Variance, } s^2 = \frac{\sum x^2 - \frac{\left(\sum x\right)^2}{n}}{n - 1}$$

The standard deviation, **s**, is simply the square root of the variance:

Standard Deviation $= \sqrt{s^2}$

The standard error is the standard deviation of the mean:-

Standard error, SE $= \sqrt{}$ (variance/sample size)

SE $= \sqrt{(s^2/n)}$

The standard error estimates how well the sample has estimated the population parameter. Standard errors are recommended for data presentation (see 6.2.2.).

The 95% confidence levels are useful as they indicate the upper and lower values between which the mean will fall 95 out of 100 times. That is, if the 95% confidence intervals of two means overlap then the means are not significantly different at $p < 0.05$. Confidence levels are calculated as:-

$$\text{mean } \pm t_{\alpha(2)} \sqrt{(s^2/n)}$$

where t is taken from the statistical tables of critical *t* values.

α is the acceptance criterion for a 2-tailed t-test, e.g. $t_{0.05}$ refers to a 5% acceptance criterion. See Sokal and Rohlf (1981) and Zar (1984, 1996) for a full description of t - tests. Section 6.3 also discusses the use of t-tests.

6.2.2 Examples

Chapter 5 details procedures for summarising the data using queries in Access from which means and variances are calculated. These summary tables can then be imported into a graphical package such as Excel to plot the data (Chapter 5 section 5.9). A simple copy and paste procedure is used to transfer a table from Access to Excel. The graphs described below provide examples of UVC and CPUE data summarised from a reef fishery in Solomon Islands which

were graphed in Excel. The data are from Samoilys et al (1995). There are three basic steps in creating charts in Excel: (i) assemble the data table; (ii) construct the chart using Chart Wizard; (iii) format the chart.

1. Catch rates (CPUE)

The mean catch rates (CPUE) are calculated in Access giving *NPUE (*number of fish per unit effort) and *WPUE* (weight of fish per unit effort) from questionnaire survey data (see Chapters 4 and 5). The data table to be copied into Excel consists of four columns from left to right: Survey; Area Code; NPUE (or WPUE) and Standard Error. The data are plotted as described in the steps below; the graphs are shown in Figure 6.1.

(i) Copy the Access data summary table and paste it into an Excel spreadsheet by highlighting the table (click on the small box separating the rows and columns on the top left of the table) and *Copy*.

(ii) To achieve the spacing in the chart insert three rows above the data for Survey 2, three rows between the data for Surveys 2 and 3, and three rows below the data for Survey 3. In the columns for number per unit effort, NPUE (a) or weight per unit effort, WPUE (b) and Standard Error fill the empty rows with zeros (0).

(iii) Highlight the data in the Area Code and NPUE columns. Choose *Insert, Chart, As New Sheet* from the menu bar, which will activate the *Chart Wizard*.

(iv) Follow the *Chart Wizard* instructions choosing the column graph because error (*SE)* bars can be plotted with this graph type. Add the appropriate axis labels, for example write *"(a) Number per unit of effort'* in the Chart Title box, write "*Survey 2 Survey 3*" in the Category (X) box, and write *"Number/Hour"* in the Value (Y) box.

(v) To add the Standard Error bars click on the columns (mean NPUE) to select the data series. Choose the *selected data series* in the Format menu, and then select Y-Error bars. Choose the *Both* display and select *Custom*. To enter the SE values click inside the +ve box and then go back to the Standard Error values in the data in the spreadsheet and highlight the relevant values. Repeat for the -ve value box. The graph will then display the SE bars.

Data analyses

(vi) To edit and format the chart, double click on the desired chart object (e.g. axes) and follow the instructions in the *Format* menu.

(vii) Repeat the above procedure for the WPUE values (Figure 6.1b).

(viii) Provide a concise but detailed caption for the graph. The caption should provide an overall title of the graph, explain what each axis represents (with units if not on axis), define any abbreviations and give the sample size (n). See Figure 6.1.

2. Stock abundance

To examine stock abundance, data from UVC surveys can be plotted to show the mean biomass of all fish surveyed (per 1000m^2) for each study site in each fishing area for three different survey times. The data used in this example come from lagoonal habitat. The data table to be copied into Excel consists of eight columns from left to right: Site; Area Code; Biomass (Survey 1); Standard Error (Survey 1); Biomass (Survey 2); Standard Error (Survey 2); Biomass (Survey 3); Standard Error (Survey 3). The data are plotted as described in the steps below; the graphs are shown in Figure 6.2.

(i) Copy the Access data summary table and paste it into an Excel spreadsheet by highlighting the table (click on the small box separating the rows and columns on the top left of the table) and *Copy*.

(ii) To achieve the spacing in the chart insert two rows above the data for Area A, two rows between the data for Areas A and B, two rows between the

data for Areas B and C, and two rows below the data for Area C. In the columns for Biomass and Standard Error fill the empty rows with zeros (0).

(iii) Highlight the data in the Area Code and Biomass (Survey 1) columns. Choose *Insert, Chart, As New Sheet* from the menu bar, which will activate the *Chart Wizard*.

(iv) Follow the *Chart Wizard* instructions choosing the column graph because error (*SE)* bars can be plotted with this graph type. Add the appropriate axis labels, for example write '*(a) Survey 1'* in the Chart Title box, insert '*A B C Area'* in the Category (X) box, and insert '*Biomass(kg/1000m^2)'* in the Value (Y) box.

(v) To add the Standard Error bars click on the columns (mean Biomass) to select the data series. Choose the *selected data series* in the *Format* menu, and then select Y-*Error bars*. Choose the *Both* display and select *Custom*. To enter the SE values click inside the +ve box and then go back to the Standard Error values in the data in the spreadsheet and highlight the relevant values. Repeat for the -ve value box. The graph will then display the SE bars.

(vi) To edit and format the chart, double click on the desired chart object (e.g. axes) and follow the instructions in the *Format* menu.

(vii) Repeat the above procedure for the Survey 2 and 3 values (Figure 6.2b &c).

Figure 6.1 Catch rates (CPUE) from questionnaire surveys in Marovo Lagoon, Solomon Islands, summarised as means and standard errors, by fishing area (A,B,C) and survey time (2,3). Catch rates represent both number of fish (NPUE) and weight of fish (WPUE) for all species combined. Sample size, n = 8. (Source: Samoilys et al 1995).

(viii) Provide a concise but detailed caption for the graph. The caption should provide an overall title of the graph, explain what each axis represents (with units if not on axis), define any abbreviations and give the sample size (n). n refers to each bar, i.e. each mean. See Figure 6.2.

The steps described above can be repeated for the stock or biomass of an individual species. The example below is for *Acanthurus nigricauda*. Exactly the same eight-step procedure described above is used to create a chart, using data summarised for one species (see Figure 6.3).

Figure 6.3 Biomass (mean total weight of all fish) of *Acanthurus nigricauda* for each site (1,2) in lagoonal habitat in each of the fishing areas (A,B,C) for the three UVC Surveys. Sample size, n = 24 for area A, and 12 for areas B and C. Error bars are standard errors. (Source: Samoilys et al 1995).

The procedures for graphing data described here constitute the first step in analysing the data. They enable the researcher to examine any apparent trends in the data. The next section describes statistical tests that are used to test whether these trends are significant.

Figure 6.2 Biomass (mean total weight of all fish) for each site (1,2) in lagoonal habitat in each of the fishing areas (A,B,C) for the three UVC Surveys. Sample size, n = 24 for area A, and 12 for areas B and C. Error bars are standard errors. (Source: Samoilys et al 1995).

Data analyses

6.3 Hypothesis testing: application of statistical tests - some basic examples

This section presents some examples of sampling designs and tests (first introduced in Chapter 2) that may be used for UVC and CPUE surveys. All the examples deal with univariate tests, but the designs can often be adapted for multivariate analyses. Note that the examples are intended as a general guide to the types of approach available and standard texts on statistics should be examined for more details.

t-test

t-tests are parametric tests used to compare two sets of data, for example the abundance of a species of fish between two reefs, or over two times on one reef. In comparing two reefs, our hypothesis might be that the abundance of coral trout on Reef 1 is different to that on Reef 2. The null hypothesis (see Chapter 2) would therefore be that the mean abundance of coral trout on Reef 1 is the same as that on Reef 2. This type of t-test is called a *2-tailed test* because we do not specify that one particular reef has more fish than another, only that they differ.

If we had hypothesised that Reef 1 had more coral trout than Reef 2, our null hypothesis would be that numbers of coral trout on Reef 1 were the same or less than those on Reef 2. This type of null hypothesis would lead to a *1-tailed test*.

The data obtained for 12 replicates at each site would be represented as follows:

Reef 1	Reef 2
$x_{1,1}$	$x_{2,1}$
$x_{1,2}$	$x_{2,2}$
$x_{1,3}$	$x_{2,3}$
.	.
.	.
.	.
$x_{1,12}$	$x_{2,12}$

where $x_{1,1}$ is the count of the first replicate at Reef 1, $x_{2,12}$ is the count of the twelfth replicate at Reef 2, etc. Computations for the t-test require the mean and variance for each reef. It is desirable but not essential that the two data sets be *balanced*, that is, have equal numbers of replicates.

Valid t-tests must satisfy the following assumptions: the data should be distributed *normally*, variances between the two data sets should be the same (i.e. *homogeneous*) and the data must be *independent*. There are tests available to determine normality and homogeneity. To test for normality, however, requires a relatively large number of replicates and some violation of this assumption is not generally considered to be a major problem in t-tests and ANOVA (Underwood 1981). Violation of homogeneity is of more concern and can lead to increased risk of Type I error (Chapter 2 section 2.4.4). A simple test of departures for homogeneity for two data sets is a 2-tailed F test, where the ratio of the largest variance to the smallest is compared. If the variances are not homogeneous, variances may be stabilised by transforming the data to a log or square root scale (Underwood 1997).

If the variances remain heterogeneous even after transformation, this is an important finding because it indicates that the distribution of that species (apart from its mean abundance) varies between the reefs compared. This finding may be considered a new observation which could lead to a new model and possibly the design of a stratified sampling program (see Chapter 2). Notwithstanding this, where assumptions of normality and homogeneity are not met, a variety of non-parametric tests (e.g. Seigel and Castellan 1988) and randomisation tests (Manly 1991) are available, which share less stringent assumptions.

To ensure independence, the replicates should be collected from randomly-selected positions within the reef at a spatial scale appropriate to the fish being sampled (see Chapter 3). Non-independence may cause variances to be smaller than they should be, which increases the chance of a Type I error (i.e. concluding there was a difference when there wasn't one). Repeated counts at exactly the same position on the reef are clearly non-independent, but there are more subtle ways in which data can be non-independent. For example, lethrinids are often very diver-wary. If replicates are obtained close together (say 10 - 20m apart), the presence of an observer at one position may affect the abundance of fish at the next position, hence the data would not be independent. It is therefore very important to design sampling programs that take into account factors such as the mobility of the species being surveyed (see also Chapter 3).

t-tests provide a good introduction to the use of parametric

statistics. They are limited, however, in that they can be applied only to two sets of data and in many situations we need to compare more than two data sets simultaneously. It is also very important to recognise that it is not valid to use multiple t-tests to compare > 2 data sets. For example, one cannot compare three reefs by using 3 separate t-tests to examine Reef 1 vs Reef 2, Reef 2 vs Reef 3 and Reef 1 vs Reef 3. This approach leads to a large increase in the risk of a Type I error (see Chapter 2). The problem is overcome by using analysis of variance (ANOVA), which allows comparisons of any number of data sets.

One-way ANOVA

ANOVA (Analysis of Variance) is a major class of parametric statistics which are highly flexible and allow us to examine very complex models. Underwood (1997) and most statistics text books provide detailed information on the theory and application of ANOVA; here only an overview is provided with some comments on how it may be applied in studies of coral reef fisheries.

For ANOVA, the size of any individual measure in a data set is assumed to be made up of three *components of variation*:

(i) the overall mean of the whole data set;

(ii) the effect of the treatment (e.g. reef, fishing pressure, etc.) that the replicate was collected from;

(iii) an individual error term associated with that particular measure.

The *error term* is also made up of three components of variation:

(a) error associated with the sampling procedure;

(b) background random variation;

(c) the variation associated with any other possible treatments we have not considered in our model.

Put differently, the *treatment* is what we consider to be the *"ecological signal"* and the *error* is the *"ecological noise"*. In a one-way ANOVA, we compare the amount of variation associated with the treatment against the variation associated with the error, which is also called the *residual*.

The model for a one way ANOVA is often shown as follows:

$$X_{ij} = \mu + A_i + e_{ij}$$

where X_{ij} is the individual measure (the jth measure in the ith set of data), μ is the overall mean of all the data, A_i is the

effect of treatment I (e.g. Reef 1 or Reef 2 or Reef 3 or Reef 4) and e_{ij} is the individual error associated with that replicate.

A typical dataset that would be analysed by a one-way ANOVA is as follows:

12 replicate counts at each of four reefs e.g.

Reef 1	Reef 2	Reef 3	Reef 4
$x_{1,1}$	$x_{2,1}$	$x_{3,1}$	$x_{4,1}$
$x_{1,2}$	$x_{2,2}$	$x_{3,2}$	$x_{4,2}$
$x_{1,3}$	$x_{2,3}$	$x_{3,3}$	$x_{4,3}$
.	.	.	.
.	.	.	.
.	.	.	.
$x_{1,12}$	$x_{2,12}$	$x_{3,12}$	$x_{4,12}$

Note that for each reef a total of 12 replicates is collected. Having the same number of replicates for each treatment provided a *balanced* statistical design. One-way ANOVA can be done with unbalanced data sets, but it is highly recommended that, wherever possible, balanced data sets be used. The ANOVA table summarising the one-way test, with r = 4 reefs and n = 12 replicates is as follows:

Source	Sums of Squares	Degrees of Freedom	Mean Square (MS)	F-ratio
Reef	SSreef	(r-1) = 3	SSreef/3	MSreefs/MSres
Residual	SSres	r(n-1) = 44	SSres/44	
Total	SStot	rn-1 = 47		

It is worth re-emphasising that when sampling reefs in the one-way ANOVA model shown above, replicates must be allocated randomly over the entire reef. Thus, if the reef covers an area of 10 hectares within the depth range of interest, the 12 replicates must be allocated randomly over that 10 ha area and not concentrated in a smaller part of the reef. By sampling over the whole area an unbiased estimate of the mean abundance (or biomass) for each reef is obtained. For example, if the mean abundance is calculated per standard area (e.g. $1000m^2$) the total abundance for the 10ha can be estimated by extrapolation (i.e. multiply by 100). If the replicates were restricted to a smaller area within the reef (e.g. a site, see Chapter 3), a biased estimate of mean population size and variance will be obtained and the total abundance calculated by extrapolation will also

Data analyses

be biased. This issue of sampling - stratification versus hierarchical sampling - is covered in Chapter 3, section 3.2.

Procedurally, there are up to three steps involved when analysing data using an ANOVA:

(i) First, treatment variances are tested for homogeneity.

Two tests are available for this - Cochran's Test which is used for balanced data sets and Bartlett's Test which can be used for balanced or unbalanced data sets. Underwood (1997) recommends using Cochran's test where possible. If variances are not homogeneous, data are usually transformed using either a log or square root transform.

(ii) Assuming homogeneity, the second step is to do the ANOVA.

If the test gives a non-significant result, we should consider whether there was likely to have been sufficient power to detect an effect (see Chapter 2, section 2.4.4). If the ANOVA is significant, we can conclude that there was a significant difference among treatments. The ANOVA does not, however, identify which treatments were significantly different. For example, all the treatments could have been significantly different, or one may have been significantly smaller or larger than the others.

(iii) To determine where significant differences lie, we use a post-ANOVA test to compare means. Generally, we use what are termed a posteriori tests.

A posteriori tests include Student - Newman - Keuls Test, Ryan's Test, etc (see Underwood 1997, Zar 1996, Day and Quinn 1989, Winer et al 1991). In some cases our hypothesis may specify a difference among levels within a treatment before the data are collected. In this case we may use *a priori* tests, which tend to have more statistical power to detect a difference among means (see Sokal and Rohlf 1981, Day and Quinn 1989).

Nested or hierarchical ANOVA

Nested ANOVA allows us to evaluate differences at different spatial and temporal scales. Generally, it works by beginning at the smallest scale considered and then working up the hierarchy to the largest scale. The test of differences in fish stocks among reefs near to a village and remote from a village (Figure 2.2 in Chapter 2) uses a hierarchical design. At its simplest level, two levels of the treatment 'reef' are

compared within each of two levels of the other treatment 'location', as shown in the following design:

Location 1		Location 2	
Reef 1	Reef 2	Reef 3	Reef 4
$x_{1,1}$	$x_{2,1}$	$x_{3,1}$	$x_{4,1}$
$x_{1,2}$	$x_{2,2}$	$x_{3,2}$	$x_{4,2}$
$x_{1,3}$	$x_{2,3}$	$x_{3,3}$	$x_{4,3}$
.	.	.	.
.	.	.	.
.	.	.	.
$x_{1,12}$	$x_{2,12}$	$x_{3,12}$	$x_{4,12}$

The model for this hierarchical design is shown as follows:

$$X_{ijk} = m + A_i + B(A)_{j(i)} + e_{ijk}$$

Here, A_i symbolises Locations and $B(A)_{j(i)}$ symbolises the Reefs nested within Locations.

Hierarchical sampling designs appropriate for nested ANOVAs are very powerful because they can be used to ***partition the variability*** among different spatial and/or temporal scales. Thus, in the above example we are interested in comparing two locations. By nesting a ***random-effects factor*** (see below) such as reefs within locations we can account for some of the natural variability within locations.

Hierarchical designs may be used to provide a valid test of the effects of different habitats on fish stocks. In the following example, we hypothesise that fish stocks occurring on slope habitats are different to those on terraces. Our main null hypothesis is that there is no difference in stocks between habitats. In order to provide a valid test of this hypothesis, however, we simply cannot sample at one site within each habitat because any differences detected may be due to the effect of some site-specific factor(s) (see section 2.3, 'confounded' effects). The most common way to overcome this is to select two or more sites within each habitat and evaluate variability between sites and then between habitats. Often we find that small-scale variability is as large (if not larger) than the hypothesised effect of habitats.

The layout of data for a comparison of habitats and sites within habitats is as follows:

Terrace habitat		Slope Habitat	
Site 1	Site 2	Site 3	Site 4
$x_{1,1}$	$x_{2,1}$	$x_{3,1}$	$x_{4,1}$
$x_{1,2}$	$x_{2,2}$	$x_{3,2}$	$x_{4,2}$
$x_{1,3}$	$x_{2,3}$	$x_{3,3}$	$x_{4,3}$
.	.	.	.
.	.	.	.
.	.	.	.
$x_{1,12}$	$x_{2,12}$	$x_{3,12}$	$x_{4,12}$

Sites 1 and 2 are in Terrace habitat, 3 and 4 in Slope habitat.

The ANOVA table summarising the hierarchical test, with h = 2 habitats, s = 2 sites within each habitat and n = 12 replicates is summarised as follows:

Source	Sums of Squares	Degrees of Freedom	Mean Square	F-ratio
Habitats	SShab	h-1 = 1	SShab/1	MShab/MSsite(hab)
Site(Hab)	SSsite(hab)	h(s-1) = 2	SSsite(hab)/2	MSsite(hab)/MSres
Residual	SSres	hs(n-1) = 44	SSres/44	-
Total	SStot	hsn-1 = 47		

Note that the residual and total degrees of freedom are the same in the hierarchical design as the one-way design and that the degrees of freedom for the reefs in the one-way design equal the total degrees of freedom for site (habitat) and habitat in the hierarchical design. The hierarchical design has partitioned the variation between habitats and sites.

The three procedures identified for the one-way analysis are the same for the nested design, including testing for homogeneity, doing the ANOVA and post-ANOVA comparisons of means if the ANOVA is significant. Note that it is important to use balanced data sets for nested designs.

Factorial and mixed model ANOVA

A factorial design is one in which the factors being compared are *orthogonal* with respect to each other. Being orthogonal means that every level of one factor occurs within every level of the other factors in the model. For example, in comparing the same two sites over two times, the factors sites and times would be orthogonal, because each time occurs within each site; and each site is surveyed at both times. This can be represented in two ways:

Time 1		Time 2	
Reef 1	Reef 2	Reef 1	Reef 2
$x_{1,1,1}$	$x_{1,2,1}$	$x_{2,1,1}$	$x_{2,2,1}$
$x_{1,1,2}$	$x_{1,2,2}$	$x_{2,1,2}$	$x_{2,2,2}$
$x_{1,1,3}$	$x_{1,2,3}$	$x_{2,1,3}$	$x_{2,2,3}$
.	.	.	.
.	.	.	.
.	.	.	.
$x_{1,1,12}$	$x_{1,2,12}$	$x_{2,1,12}$	$x_{2,2,12}$

or:

Reef 1		Reef 2	
Time 1	Time 2	Time 1	Time 2
$x_{1,1,1}$	$x_{1,2,1}$	$x_{2,1,1}$	$x_{2,2,1}$
$x_{1,1,2}$	$x_{1,2,2}$	$x_{2,1,2}$	$x_{2,2,2}$
$x_{1,1,3}$	$x_{1,2,3}$	$x_{2,1,3}$	$x_{2,2,3}$
.	.	.	.
.	.	.	.
.	.	.	.
$x_{1,1,12}$	$x_{1,2,12}$	$x_{2,1,12}$	$x_{2,2,12}$

The main distinction between orthogonal and hierarchical factors is that for the latter, each level of each factor is not represented in the levels of the other factors. This is illustrated by the examples provided for the hierarchical design. For example, it is impossible by definition for Site 4 to occur in the Terrace Habitat, or for Site 1 to occur in the Slope Habitat; hence these sites are nested within the respective habitats. Another distinction between orthogonal and hierarchical designs is that orthogonal designs have *interactive terms*. An interaction may occur where variation among levels of one factor is inconsistent among levels of the other(s). The interaction for a simple two-way orthogonal test is shown as follows:

$$X_{ijk} = m + A_i + B_j + AB_{ij} + e_{ijk}$$

Here, A_i and B_j are the *main effects* of times and reefs and AB_{ij} is the *interaction term*. It is important to note that where a significant interaction is found, the results cannot

Data analyses

be interpreted in terms of the main effects. This is because the significant interaction means that the effect on each factor is dependent upon the levels of the other factor(s). A detailed explanation of factorial designs and their interactions is beyond the scope of this manual. Fisheries scientists seeking to use these designs should obtain training in this area. Useful texts include Underwood (1997) and Winer et al (1991), which include worked examples.

Another concept that is important to consider in ANOVA is whether factors are to be treated as *fixed* or *random* factors. This has major consequences for the way in which the analysis is done and how the results are interpreted and applied. The levels compared for fixed factors generally consist of only a few specific cases that could be compared, whereas random factors contain levels drawn from a large array of potential levels. The following examples illustrate this distinction. In a design comparing the factor fished vs unfished reefs, the condition of fishing would be a fixed factor because the reefs are either fished or unfished (not some other condition). If the design were extended to include, say, three levels of fishing such as high, medium and low pressure, these three levels would also constitute a fixed factor.

If we were interested in comparing abundance of fish on several specific reefs (because we may be concerned about levels of fishing there), the reefs would be a fixed factor. On the other hand, if we wished to know something about fish stocks generally within a region, we might randomly select reefs from a large number that are potentially available. i.e. only a few reefs are included. In this case reefs would be a random factor. Factors which are fixed limit the generality of how we interpret the results: in our example of reefs being a fixed factor, we may draw conclusions only about those reefs. Random factors are more generally applicable and we may draw conclusions from our test about the region from which the reefs were randomly selected.

The notion of fixed versus random factors is complex and is, like the discussion of interactions, beyond the scope of the manual. Fisheries officers needing to apply these tests must however, be aware of the implications of any design they intend to use. Moreover, statistical computer programs generally require that fixed or random and nested or orthogonal factors are specified. Where these cannot be, or are not specified, the typical default is for the computer program to consider factors as being fixed and orthogonal.

If unsure, it is important to seek advice and/or to run some data in which the outcome for an appropriate design is known (e.g. from one of the texts such as Winer et al 1991).

The following table shows the design for the two-way orthogonal design discussed in this section. It can be seen that the formulae vary for calculation of the F-ratio for the main effects, depending on whether factors are fixed or random.

Source	Sums of Squares	Degrees of Freedom	Mean Square	F-ratio denominator (MS)*			
				1	2	3	4
Time	SStime	t-1 = 1	SStime/1	res	t x s	res	t x s
Site	SSsite	s-1 = 1	SSsite/1	res	res	t x s	t x s
T x S	SSt x s	(t-1)(s-1) = 1	SSt x s/1	res	res	res	res
Residual	SSres	ts(n-1) = 44	SSres/44	-	-	-	-
Total	SStot	tsn-1 = 47					

* 1 = Time, Site fixed; 2 = Time fixed, Site random; 3 = Time random, Site fixed; 4 = Time, Site random.

One of the important features of ANOVA is that it is possible to combine orthogonal and hierarchical factors within the same analytical framework. Thus, in a survey which compares reefs within different habitats over time, habitats and times will be orthogonal to each other; times and reefs will be orthogonal; but reefs will be nested within habitats. This design will yield two interaction terms: times x habitats and times x reefs (habitats), but there cannot be an interaction between reefs and habitats. These mixed designs are quite common in ecology and fisheries officers required to evaluate complex questions will often need to utilise such designs.

Asymmetrical ANOVA

In the past decade there has been extensive work on how to detect the effects of human activities on the aquatic environment. Much of this work has required the development of new kinds of analytical designs to test hypotheses. One class of designs now being used in environmental impact assessment involves the use of asymmetrical ANOVA which allows us to compare a single impact site against two or more reference sites. Essentially,

this provides a means of comparing variability between the impact site with that occurring among reference sites. Recently, this approach has been adopted for exploited invertebrates on coral reefs in Solomon Islands (Lincoln Smith and Bell 1996) and should be applicable to other tropical fisheries. A good description of asymmetrical ANOVA is provided by Underwood (1997).

Regression and ANCOVA

Regression and analysis of covariance (ANCOVA) are important statistical tools in fisheries science and they are used in variety of situations, such as when examining the relationship between fishing effort and yield. Broadly they are described as follows.

Regression, like ANOVA, is one of the General Linear Models (Sokal and Rohlf 1981, Zar 1984). It allows us to examine the relationship between one or more independent variables and a dependent variable. For example, in comparing the size of fish stocks on reefs with fishing pressure on those reefs, the size of the stock would be the dependent variable and the amount of fishing effort would be considered to be the independent variable. If the relationship is significant, the regression line can then be used to predict how the dependent variable will vary as the independent variable changes. The relationship is expressed as a r^2 value, termed the **coefficient of determination** (Zar 1984). Either t-tests or ANOVA procedures may be used to test the significance of the relationship between the two variables. As a rule of thumb the r^2 value, converted to a percentage, represents the amount of variance in the data that is described by the regression model. It is possible to have a large r^2 value (e.g. .90 or 90%), but because of, say, small sample size the relationship is not significant. Conversely, one may have a small r^2 value but a significant relationship. Plotting the data will illustrate these scenarios.

Two extensions of simple regressions are multiple regressions and ANCOVA. Multiple regression examines how a dependent variable changes in relation to two or more independent variables. For example, we may hypothesize that fish stocks vary on reefs according to fishing pressure and percentage cover of live coral. By measuring the abundance of fish, the amount of fishing and the coral cover on a number of reefs, we can examine the relationship between these factors. Multiple regression allows us to test if there is a relationship and to measure the contribution of each independent variable to the relationship.

ANCOVA provides a means of comparing two or more regressions. For example, we may have data on the relationship between fishing pressure and stock size from several reefs in two regions. We can use ANCOVA to determine firstly if the slopes of the regressions are statistically similar and if they are to then determine if the elevation of the regression lines is significantly different.

There are several important issues that need to be considered when using regression and ANCOVA. First, relationships between variables may be linear or non-linear and most regression models assume linearity. It is important to plot the data to see if the relationship appears to be linear. Sokal and Rohlf (1981) provide a means of testing for linearity. Second, valid regression and ANCOVA must satisfy the assumptions of normality, homogeneity and independence of data. Moreover, if variances are found to be heterogeneous, transforming the data can cause the relationship between the variables to change.

Chi-Square and concordance tests

Tests based on goodness-of-fit or ranks may be very useful in studies of fisheries, where the assumptions of parametric tests cannot be met and where data sets are very unbalanced. These situations can often occur in creel and questionnaire surveys. The assumption of independence of data must, however, still be met. Here two tests are considered, chi square tests and tests of concordance. Readers should refer to an excellent statistical text: Siegel and Castellan (1988), which provides details on methods and statistical tables required.

Chi square tests are generally used to compare two or more groups with respect to some characteristic and, therefore, with respect to the relative frequency with which group members fall in two or more categories (Siegel and Castellan 1988). The test is based on the notion of comparing the observed condition with what might be expected either by chance alone or derived from some independent information.

For example, we may need to determine the likelihood that similar proportions of male and female fish are taken by an artisanal fishery. Without any prior information, it is likely that we would compare the observed proportion of males and females in the catch with an expected proportion of 0.5. However, we may know from prior information that the species changes sex as they grow and that those fish entering the fishery are most likely to be females. In this

case, the expected proportion is likely to be much larger for females (say, 0.75). Finding a disproportionately large or small proportion of females in the catch may lead to a hypothesis that the fishery is having an adverse effect on the population structure of the species.

Another application of chi square tests is with respect to the relative contribution of a group of species within an artisanal fishery. For example, a chi square test could be used to examine whether the proportion of lutjanids sold in local markets varied from one village to another. A test of this hypothesis would be to inspect the market catch at several villages and determine the relative abundance of lutjanids being displayed for sale. This approach may be far easier to undertake than to design a balanced test with replication.

Tests of concordance are very useful for dealing with ranked data and can be widely applied in analysis of fisheries questionnaires. One useful test is Kendall's Coefficient of Concordance (*W*), which is presented in Siegel and Castellan (1988). A similar test, known as Anderson's Test, is described in Winer et al (1991).

An example of the application of a concordance test would be in seeking to determine which are the most productive reefs for fishing in the vicinity of a village. Here we might ask 10 of the most experienced fishermen (or women) to rank 12 reefs of interest in terms of fish catch. The test of concordance examines whether there is general agreement among the fishers about how they would rank fishing reefs. Another example might be to ask the same fishers which season produces their largest catches. Testing these types of questions can provide important information about how communities view their fisheries and the extent to which there is local agreement within the communities. They can also be very important as part of pilot or frame studies, in helping scientists to refine their study designs. For example, finding large concordance among fishers with respect to popular fishing reefs may suggest hypotheses that can be investigated and may help with selection of study sites.

6.4   Potential yields

This section addresses the third step outlined in the introduction to this chapter: *stock assessment*. In this section we describe the application of UVC estimates of stock abundance to calculating yields, as a means of illustrating some basic stock assessment concepts. As stated earlier,

this manual does not cover the procedures for using stock assessment models such as surplus production models or yield per recruit models.

Yields, usually expressed as tons per square kilometre per year (t km^{-2} yr^{-1}) are the most widely used measure of a reef fishery and provide a useful basis for comparisons with other areas or fisheries. Reef fishery yields around the world have been reviewed by Marshall (1980), Russ (1984), Munro and Williams (1985), Russ (1991) and Dalzell (1996). An average sustainable yield of 5 t km^{-2} yr^{-1} for coral reefs is often cited, though measured yields vary widely. For example, yields on fringing reefs in the Pacific range from as low as 0.3 t km^{-2} yr^{-1} in Fiji (Jennings and Polunin 1995) to 14.0 t km^{-2} yr^{-1} in the Philippines (Alcala and Gomez 1985, Dalzell 1996).

When using yields to describe a fishery there are two important factors to consider:

(i) what species are involved?

(ii) what area do the data relate to?

Yields for shallow-water reef fisheries have been reported in a number of studies, but the suite of species may vary considerably (Russ 1991, Dalzell 1996). For example, the inclusion or exclusion of the generally larger pelagic or semi-pelagic species such as Carangidae, Scombridae or sharks will have a considerable effect on the size of the yield. Clearly, it is important to specify which species are included in the yield estimates, and to consider this factor when comparing yields between different areas or studies.

Similarly, the area the yields refer to, whether it is strictly coralline habitat, or whether it includes sand, seagrass, and mangrove habitats, or whether it extends to depths beyond the shallow reef slope, will make a considerable difference to the size of the yield (Marshall 1980, Russ 1984, Munro and Williams 1985, Russ 1991 and Dalzell 1996). For example, Jennings and Polunin (1995) demonstrate that average yields calculated for strictly coralline areas in Fiji are around 10.2 t km^{-2} yr^{-1} compared with 3.4 t km^{-2} yr^{-1} for the overall fishing grounds in the same region.

A standard method for estimating yields from fishery-dependent catch data involves scaling up (extrapolating) catch records e.g. from questionnaire surveys, creel surveys (Chapter 4) or logbooks, to the total area of the fishing ground. The latter is usually calculated from charts and aerial photographs. Jennings and Polunin (1995) provide a

clearly worked example of this approach for reef fisheries in Fiji.

Simple calculations for estimating yields from UVC data are detailed below to illustrate the potential information that may be obtained from fishery-independent surveys such as UVC surveys. Yields are predicted for two reefs in both Australia and Fiji (data from Samoilys and Carlos 1992) using hypothetical estimates of fishing mortality based on two different exploitation regimes, "heavy" and "light" fishing pressure. ***The predicted yields are not estimates of sustainable yield***. They serve simply to demonstrate how independent estimates of stock biomass can be used in a surplus production model framework (see Chapter 1, section 1.2.3) for a first approximation of potential yield.

The calculations use some of the most basic equations of fish population dynamics (see standard fisheries texts, e.g. Gulland 1971, Ricker 1975, Gulland 1983, Sparre and Venema 1992) which relate yield, exploitation rates (mortalities) and biomass:-

$$Y = F \times \overline{B} \qquad (1)$$
Y = annual yield
F = fishing mortality
$\overline{B}$ = average biomass

Since fishing mortality is not known it may be derived from the following equation:

Exploitation Rate, E = Fishing Mortality, F / Total Mortality, Z

Since Z = F + M

where M = natural mortality, we can write:

$$E = \frac{F}{F+M} \qquad (2)$$

Using equation **(2)**, fishing mortalities were calculated from known natural mortalities and two hypothetical exploitation rates. The derived fishing mortalities were applied to equation **(1)** to calculate predicted yields of stocks in different theoretical exploitation regimes, using UVC measures of population biomass.

An example of the computations is given for biomass estimates of Roving Serranidae obtained by Samoilys and Carlos (1992) in Fiji and Australia (Table 6.1). The Roving Serranidae are predominantly the coral trout (*Plectropomus* spp.), which are the most important species in the Great Barrier Reef fisheries. An estimate of natural mortality for

Plectropomus leopardus of 0.46 was obtained from Fishbase, ICLARM's CD-ROM biological database on fish (Fishbase 1995). The computations for the hypothetical estimates of fishing mortality are as follows:

Natural mortality of *Plectropomus leopardus* = 0.46

(i) Light exploitation rate, E = 0.1 (10%)

(ii) Heavy exploitation rate, E = 0.5 (50%)

These exploitation rates can be viewed in the context of maximum sustainable yields, MSY. Gulland (1971) proposed that MSY is obtained when E = 0.5, i.e. when fishing mortality = natural mortality (equation (2) above). Others believe this overestimates MSY, and in fact E_{opt} is closer to 0.2, i.e. fishing mortality is approximately 1/3 of natural mortality (Beddington and Cooke 1983). Here we use the equivalent of Gulland's estimator to represent heavy fishing pressure, and 1/2 the MSY estimator proposed by Beddington and Cooke to represent low fishing pressure.

(i) Light exploitation rate

$$E = \frac{F}{F+M} \qquad (2)$$

$$0.1 = \frac{F}{F+0.46}$$

$$F = 0.051$$

(ii) Heavy exploitation rate

$$0.5 = \frac{F}{F+0.46}$$

$$F = 0.46$$

These fishing mortalities are then used in equation **(1)** to derive predicted yields (Table 6.1).

The calculations show that for each pair of reefs the predicted annual yield is distinctly different. A high exploitation rate (E=0.5) on the GBR would yield almost 3 times as much catch on Escape Reef compared to Scott Reef. These differences can be interpreted in various ways. The contrast in predicted yield between the two reefs could be due to differences in the ecological productivity of the reefs. Alternatively, it could relate to the fact that the reefs are currently and/or were historically subjected to very different fishing mortalities. The calculations serve to

Data analyses

Table 6.1 Calculations of predicted yields, **Y,** from reefs in Australia and Fiji, using the catch equation **(1)**, biomass estimates, $\overline{B}$, from UVC surveys and two fishing mortalities, **F** (modified from Samoilys and Carlos 1992).

ESCAPE REEF (Australia)	**SCOTT REEF** (Australia)
Roving serranid biomass	Roving serranid biomass
$= 2.75 \, kg/1000m^2$	$= 0.96 \, kg/1000m^2$
$Y = F.\overline{B}$ (1)	$Y = F.\overline{B}$ (1)
(i) light fishing:	(i) light fishing:
$Y = 0.051 \times 2.75$	$Y = 0.051 \times 0.96$
$= 0.140 \, kg/1000m^2/yr$	$= 0.049 \, kg/1000m^2/yr$
$= \mathbf{140 \, kg \, km^{-2} \, yr^{-1}}$	$= \mathbf{49 \, kg \, km^{-2} \, yr^{-1}}$
(ii) heavy fishing:	(ii) heavy fishing:
$Y = 0.46 \times 2.75$	$Y = 0.46 \times 0.96$
$= 1.265 \, kg/1000m^2/yr$	$= 0.442 \, kg/1000m^2/yr$
$= \mathbf{1,265 \, kg \, km^{-2} \, yr^{-1}}$	$= \mathbf{442 \, kg \, km^{-2} \, yr^{-1}}$
NAITAUBA REEF (Fiji)	**DUFF REEF** (Fiji)
Roving serranid biomass	Roving Serranid biomass
$= 0.94 \, kg/1000m^2$	$= 0.17 \, kg/1000m^2$
$Y = F.\overline{B}$ (1)	$Y = F.\overline{B}$ (1)
(i) light fishing:	(i) light fishing:
$Y = 0.051 \times 0.94$	$Y = 0.051 \times 0.17$
$= 0.048 \, kg/1000m^2/yr$	$= 0.009 \, kg/1000m^2/yr$
$= \mathbf{48 \, kg \, km^{-2} \, yr^{-1}}$	$= \mathbf{9 \, kg \, km^{-2} \, yr^{-1}}$
(ii) heavy fishing:	(ii) heavy fishing:
$Y = 0.46 \times 0.94$	$Y = 0.46 \times 0.17$
$= 0.432 \, kg/1000m^2/yr$	$= 0.078 \, kg/1000m^2/yr$
$= \mathbf{432 \, kg \, km^{-2} \, yr^{-1}}$	$= \mathbf{78 \, kg \, km^{-2} \, yr^{-1}}$

illustrate the relative differences between reefs based on a single estimate of stock biomass. *It is important to understand that we cannot specify what is the maximum sustainable yield (MSY) from these calculations*.

To calculate MSY the generalised model proposed by Garcia et al (1989) is recommended (see Chapter 1, Appeldoorn 1996):-

$$MSY = \frac{BM^2}{2M - F}$$

Where $\overline{B}$ is the average exploited biomass

M is the natural mortality

F is the fishing mortality

Using the same Biomass and Mortality estimates (high exploitation rate) from Escape Reef (Table 6.1) we can derive MSY as follows:

$$MSY = \frac{2.75 \times (0.46)^2}{2(0.46) - 0.46}$$

$MSY = 1.265 \, kg \, 1000m^{-2} \, yr^{-1}$

$= 1,265 \, kg \, km^{-2} \, yr^{-1}$

In this example the MSY value is hypothetical since F was not known and was based on a "high" exploitation rate of 0.5. It therefore gives us the same value as the predicted yield calculated previously (Table 6.1).

CHAPTER SEVEN: REPORTING

Krishna Swamy

7.1 Introduction

In many Pacific Island countries research carried out by fisheries scientists is often not documented, and therefore the results are not communicated to the broader community - the scientific community, other interest groups such as resource custodians, fishers and the general public. Although in some cases data may be stored, accessing this information may be difficult and time consuming. In many cases the main aim of the research was to communicate the results and recommendations to managers. The research is often considered completed once the managers are satisfied and recommendations have been adopted. In addition, research methodology, and the generation of data has often been considered secondary by fisheries scientists in Pacific Island countries.

In fisheries stock assessment work, historical data are very important, therefore, documentation and storage of research data is vital. Fisheries departments in general have a history of poor database management. This has resulted in the loss of valuable data and research being repeated. Chapter 5 describes methods for designing and managing databases so that data stored can continue to be accessed and maintained over time.

A fisheries researcher should communicate his/her results to as wide an audience as possible, otherwise the work has diminished value (Stapleton, 1987). Many fisheries researchers from the Pacific region are not able to express themselves properly in English, since it is often their second language. To overcome this problem, Stapleton (1987) suggests writing the first draft of the report in the vernacular language and later translating it into English. This Chapter provides guidelines on producing technical reports and for communicating research results to managers and the broader community through various media.

7.2 Technical reports

A technical report details all aspects of a given research project, not just the results and recommendations. The preparation of a technical report takes time and effort and it is important to factor this in to one's research program.

As stated above, the research is of limited value if it is not documented, and the most effective form of documentation is a technical report or publication (see below). The main audiences for technical reports are the resource managers and other researchers. Although technical reports are sometimes bulky, it is essential to include the results, summarised either graphically or in tables. This ensures the information collected is presented and available to other researchers. This is especially important for future studies that may wish to compare results with a previous study and thereby produce a temporal or historic analysis of data.

This section does not cover every aspect of report writing, but summarises the main points. It is expected that this chapter will be most useful to fisheries scientists who are not native English speakers and who have little experience in writing technical reports. The structure and the layout of technical reports may vary; the basic components and the most common layout are discussed below. One of the most effective ways of learning how to write good technical reports or papers is to read those produced by well recognised researchers in the same field, to experience content and writing techniques. Seeking reviews from peers (see section 7.5 below) is also very helpful and highly recommended.

7.2.1 Executive summary

The executive summary is typically designed for resource managers as they do not have the time and/or the technical understanding to read the whole report. Therefore, this section should be confined to three or fewer pages. Other researchers will also read the summary to find out if the report is relevant to their work. Therefore, this section must summarise the most significant results from the research. The executive summary should briefly cover each of the sections of the report, and include the important issues that have been discussed in the results and discussion sections.

7.2.2 Introduction

The introduction gives a brief description of the problem or question the research project is addressing and clearly states the objectives of the study. Chapter 2 describes

procedures for identifying the questions and objectives of a research project. These are documented in the introduction of the technical report, which also covers background information such as previous research done in the area or on the topic. Sometimes secondary objectives may be set for the project; these should also be mentioned. Even if the objectives were not all achieved, they should still be documented. This is especially relevant to funding agents who have provided funds to address the objectives presented in the project proposal.

7.2.3 Materials and methods

This section provides a clear description of the materials and equipment used in the research work. The methodology used must also be described carefully, and the time and duration of the work should be mentioned. A simple rule for writing the methods section is to remember that it should be possible to repeat the research by following the description of the methodology. This section is especially important for comparisons with similar studies and for interpreting the results. The statistical analyses employed should also be described, usually in a sub-section entitled "data analyses", which would include any data transformations and preliminary tests for normality, homogeneity of variances, etc (see Chapter 6) that were used.

7.2.4 Results

This section describes the results of the research and should be compiled carefully, succinctly and without discussing the results. Wherever possible, the results should be presented in figures and tables supported by explanatory text. Mean values should be given together with some measure of variation such as standard error or 95% confidence limits (see Chapter 6). The results of statistical analyses should be also presented and clearly explained, though not discussed. As a rule of thumb this section does not cite other literature.

7.2.5 Discussion

This section interprets and discusses the results and their implications. It is important to compare the results with other (preferably published) work. It is also advisable to discuss any problems or drawbacks in the research and suggest ways of improving the quality of the results in future studies. The discussion refers back to the introduction to put the results in the context of the objectives of the research and background information, and cites other literature on the subject.

7.2.6 Conclusion - recommendations

Not all reports will include a conclusion and recommendations section, but it is useful to outline the main findings of the results and their implications for resource managers. The implications can also lead to recommendations. The resource managers will want to know what could be done and how the recommendations could be implemented. It is helpful to prioritise the recommendations and list them accordingly. This section requires careful thought since the recommendations must be achievable and realistic to resource owners and users. Some recommendations may initially be unpopular to fishers or custodians; if so, other forms of communication are probably vital if they are to be implemented. This could be one of the recommendations in the technical report: that the recommendation is explained to the local community using other media (see section 7.7).

7.2.7 Acknowledgments

It is important to recognise all those who have helped in the research by acknowledging them in this section, usually placed in the beginning of the report. All significant funding contributions should be mentioned, together with any technical or practical support, such as field work (e.g. boat crews), lab work and the preparation of figures (technicians), and typing.

7.2.8 References

All the literature that has been cited in the report must be listed in full in the references section at the back of the report. There are several standard formats for compiling references which are used by various scientific journals. It is important to be consistent and use the same format throughout the reference list. When preparing a paper for publication one must follow the format requested by the journal.

7.2.9 Appendices

Appendices are useful for presenting additional reference information that is not closely tied to the text of the report. For example, large tables of results, figures, graphs, notes

on statistics or software packages etc, may all be placed in appendices. This avoids breaking the reader's concentration, and also supplies additional relevant information for those interested. As an example, it is useful to provide the mean density and biomass (with standard errors) per species for each UVC survey (Samoilys et al 1995), and it would be inappropriate to put this in the main body of the report. Instead, the results section of the report would present this data in a more summarised form, for example graphically by family.

7.2.10 Nomenclature

Often common, scientific and vernacular names are incorrectly presented in technical reports. All scientific names must either be underlined or in italics. The use of italics or underlining must then be consistent through the report. It is acceptable to use local or common names as long as the scientific name is provided when it first appears in the report. In addition, it is important to carefully check the scientific names with the latest taxonomic references, and provide the reference source. Scientific names of coral reef fishes change frequently due to ongoing taxonomic research.

7.2.11 Tables, figures and graphs

Figures and tables can present valuable information with great clarity if presented properly. They can convey relatively complex results at a glance, and therefore must be included in technical reports and scientific publications. Every table or figure presented must be discussed or referred to in the report. All tables and figures must be numbered, and this is usually done in sequence or with reference to section numbers in the report. The numbering method is a personal choice; the simplest method is recommended for easy cross referencing. Most word processors can automatically create tables of contents even for tables and figures. Sequential numbering is advantageous, as most word processors are able to update the numbering both within the text and in the table of contents.

The naming of figures and tables requires careful consideration (see Chapter 6 Section 6.2.2). The caption (title) should explain clearly and concisely what the table or figure represents without having to refer to the main text of the report. The caption should not be duplicated in the text of the report; in other words the caption should not be used as an explanation of the figure or table. Instead, the text of the report should expand on the caption. Captions should appear above the table, and beneath the figure.

7.3 Data storage

It is not possible or advisable to include all the raw data in a technical report. Sometimes it is advisable to place some of the raw data in the actual report as an appendix, depending on the requirements (e.g. of the funding agent). Raw data are stored separately, both in electronic form and as hard copy. Chapter 5 discusses database management including the archiving of data, and the development of a standard operating procedure (SOP) which ensures there are clear instructions for data retrieval. The storage of raw data is vital for future researchers and for further work. This is an important issue in the region since the same data may be highly relevant to fisheries scientists in other countries. Most Pacific Island Fisheries Divisions do not have facilities to store data that are easily accessible to other users, therefore it is advisable to store a copy of the data with a regional organisation. The most suitable organisations are: the South Pacific Commission (SPC) and the Forum Fisheries Agency (FFA), with the former probably being more appropriate since SPC also carries out research work.

7.4 Publication

The circulation of technical reports is limited, therefore every effort should be made to publish for a wider international audience. It is important as a fisheries scientist to communicate one's research findings to other workers in the same field of study. This will stimulate further discussion and research on the subject. After the completion of a technical report one should prepare a manuscript for publication, preferably writing both concurrently. Fisheries scientists are generally assigned to a new project after the completion of a research study. Consequently it is difficult to return to the previous project to write papers. In addition, over time one tends to forget some details of the work that has been completed, lose some of the information collected or misplace some of the notes compiled earlier. Therefore it is strongly recommended that manuscripts for publication are prepared while writing the technical report.

Choosing the right journal for a scientific paper is based on familiarity with the journals and the literature. Seeking advice from other researchers who are experienced in publishing is recommended. Although the basic format of a publication manuscript will be similar to a technical report,

Reporting

every journal has its own format and style which one must follow. Recent papers in the journal should be used as a reference for format and style and a style guide can be requested from the journal.

7.5 Scientific reviewers

It is highly advisable to have the draft report or paper reviewed by experts from the same field. For scientists with limited exposure to writing technical reports and scientific papers, any comments and suggestions on the report or paper will be valuable. With the scarcity of fisheries scientists in the Pacific Island region, liaising with scientists from the South Pacific Commission and from Australia is strongly recommended. Most scientists working in the same field of research will be happy to review and comment on a manuscript.

7.6 Communicating results to managers

When communicating with fisheries managers from Pacific Island countries, one must realise that most of them have little knowledge of the principles of stock assessment. Therefore, any fisheries stock assessment report should have a section dedicated to resource managers. As discussed above (section 7.2.1), this section is the executive summary in a technical report. It is useful for managers if the research work is presented point by point in the executive summary. This would include the recommendations (see section 7.2.6), which are usually presented in order of importance. Although fisheries technical reports may be read by a wide audience, recommendations are usually designed for resource managers and resource custodians. Therefore, the recommendations must be precise, applicable and achievable in terms of available resources (e.g. funds and people). The recommendations should be phrased in such a way that they can be easily incorporated into the reports or submissions that the resource managers write. This will also result in faster action in response to the recommendations.

7.7 Communicating results to the resource custodians, fishers and the community

Many resource custodians and fishers are realising that their resources are limited. In countries where a customary fishing rights system is strong, the dissemination of information to local custodians is important. In many areas custodians take a very active role in managing their resources using traditional management methods. However, these are not always effective where modern fishing gear is used extensively. It is advisable to inform resource custodians of the basics of population dynamics and stock assessment. Once these principles are understood resource owners are more likely to appreciate the research work and there is a greater chance that the recommendations will be accepted. As mentioned above, there is no point in presenting recommendations that are not feasible because they simply will not be adopted. However, recommendations that may appear initially unpopular, should be carefully communicated (see below) if custodians and fishers are to accept them.

There are several ways in which information can be transmitted to the public and the resource custodians. Given the remoteness of coastal villages and islands scattered over a wide area, communication to these villages is a major problem. The simplest and the easiest form of publicity is the local radio broadcast. Pamphlets, videos and newspaper articles are also effective communication tools. Public meetings in villages in which fisheries officers present and discuss the research, though obviously more expensive to conduct, are effective and popular and provide an opportunity for feedback from the local community.

7.7.1 Video/television

The introduction of television and videos has proved very effective in transmitting information to those Pacific Islands where television is easily accessible. Video production can be expensive and time consuming, but if resources are available, it is well worth the cost since one is able to reach a much wider section of the community compared with technical reports. Video has proved to be a very effective means of disseminating information quickly, and is strongly recommended for communicating with resource custodians, fishers and local communities. Wherever possible, the material should be produced in the vernacular language. When producing videos special attention should be given to school children. If the information is understood by children, it will slowly be relayed to the elders responsible for managing the fisheries resources.

7.7.2 Posters and pamphlets

Every effort should be made to produce information leaflets for resource custodians and fishers. The purpose of leaflets

and posters is mainly to disseminate basic information both in graphic and written forms to a wide cross section of the community. Leaflets and posters are an effective communication medium because copies are easily available to anyone interested in the subject, at a minimum cost. Furthermore, they can be displayed for longer periods than video or television. In the Pacific region, leaflets and posters are seldom reprinted, thus this informative material is often not available after the completion of a project. It is therefore advisable to make provisions to reprint this material. Posters and leaflets should be distributed to custodians, fishers and other interested members of the local community. Schools within the area should also be included in the distribution list.

7.7.3 Radio broadcasts

Radio broadcast is the simplest and the cheapest (sometimes free) method for communicating information to custodians, fishers and the general public. Radio broadcasts have a wide reach and therefore the dissemination of information is very effective. Since it is not possible to choose a specialist audience, it is only possible to broadcast general information. Thus, radio is a good tool for educating the general public. Tips on fishery matters are usually well received by the general public as most of them are affected by the sea in some way or other. Broadcasting in vernacular languages is most effective because the information can be easily understood; however producing English versions can also be beneficial.

7.7.4 Newspapers

Newspapers are widely read in urban areas. Most urban centres are now populated by people from rural areas. In recent times the public has become vocal in regard to fisheries resources. This is evident from the number of letters regarding fisheries matters being printed in the daily newspapers. Custodians who have access to newspapers are using this medium to highlight problems regarding their fisheries resources. These custodians could be targeted by publishing articles in daily newspapers. Newspaper articles should be fairly general in nature and geared towards educating the public. Furthermore, newspaper articles are archived by libraries, and will therefore serve as a historic reference source for future users.

Reporting

REFERENCES

Alcala, A.C., Gomez, E.D. (1985). Fish yields of coral reefs in the central Philippines. Proc. 5th Int. Coral Reef Symp. Tahiti 5:521-524

Andrew, N.L., Mapstone, B.D. (1987). Sampling and the description of spatial pattern in marine ecology. Oceanogr. Mar. Biol. Ann. Rev. 25:39-90

Appeldoorn, R. (1996). Model and method in reef fishery assessment. pp. 219-248. In: Polunin, N.V.C., Roberts, C.M., (eds.) Reef Fisheries, Chapman & Hall, London

Ayling, A.M. & Ayling, A.L. (1992). Bramble Reef replenishment area - baseline survey. Report to the Great Barrier Reef Marine Park Authority. 22pp

Bainbridge, S.J., Baker, V.J. (1994). Database Design and Operation. pp. 313-328. In: Survey Manual for Tropical Marine Resources, English, S., Wilkinson, C., Baker, V., (eds.) ASEAN-Australia Marine Science Project, Australian Institute of Marine Science, Townsville, Queensland

Bazigos, G.P. (1974). The design of fisheries statistical surveys - inland waters. FAO Fisheries Technical Paper 133. 122pp

Beamish, R.J., McFarlane, G.A. (1983). The forgotten requirement for age validation in fisheries biology. Trans. Am. Fish. Soc. 112(6): 735-743

Beddington, J.R., Cooke, J.C. (1983). The potential yield of fish stocks. FAO Fisheries Technical Paper, 242. 47pp

Bell, J.D., Craik, G.J.S., Pollard, D.A., Russell, B.C. (1985). Estimating length frequency distributions of large reef fish underwater. Coral Reefs 4:41-44

Bohnsack, J.A., Bannerot, S.P. (1986). A stationary visual census technique for quantitatively assessing community structure of coral reef fishes. NOAA Technical Report NMFS 41:1-15

Brander, K. (1975). Guidelines for collection and compilation of fisheries statistics. FAO Fisheries Technical Paper, 148. 46pp

Brock, V.E. (1954). A preliminary report on a method of estimating reef fish populations. Journal of Wildlife Management 18:297-308

Bros, W.E., Cowell, B.C. (1987). A technique for optimizing sample size (replication). J. of Exper. Mar. Biol. and Ecol. 114:63-71

Caddy, J.F., Bazigos, G.P. (1985). Practical guidelines for statistical monitoring of fisheries in manpower limited situations. FAO Fisheries Technical Paper, 257. 86pp

Caddy, J.F., Mahon, R. (1995). Reference points for fisheries management. FAO Technical Paper, 347. 83pp

Cappo, M., Brown, I.W. (1996). Evaluation of sampling methods for reef fish populations of commercial and recreational interest. CRC Reef research Centre, Tech. Rep. No.6, Townsville, 72pp

Carpenter, S.R. (1989). Large-scale perturbations: opportunities for innovation. Ecology 71:2038-2043

Clarke, K.R. (1993). Non-parametric multivariate analysis of changes in community structure. Australian Journal of Ecology 18:117-143

Cohen, J. (1988). Statistical Power Analysis for Behavioral Sciences. Lawrence Ehrlbaum and Associates, New Jersey

Dalzell, P. (1996). Catch rates, selectivity and yields of reef fishing. pp. 161-192. In: Polunin, N.V.C., Roberts, C.M., (eds.) Reef Fisheries, 1st edn., Chapman and Hall, London

Day, R.W., Quinn, G.P. (1989). Comparison of treatments after an analysis of variance in ecology. Ecol. Monog. 59:433-463

Doherty, P. (1991). Spatial and Temporal Patterns in Recruitment. pp. 261-293, In: Sale, P.F., (ed.) The Ecology of Fishes on Coral Reefs, Academic Press, Inc., San Diego

Doherty, P.J., Fowler, T. (1994). An empirical test of recruitment limitation in a coral reef fish. Science 263:935-939

Doherty, P.J., McB. Williams, D. (1988). The replenishment of reef fish populations. Oceanog. Mar. Biol. 26:487-551

Doherty, P.J., Fowler, A.J., Samoilys, M.A., Harris, D.A. (1994). Monitoring the replenishment of coral trout (Pisces:Serranidae) populations. Bull. Mar. Sci. 54(1):343-355

Doubleday, W.G., Rivard, D. (1983). Sampling commercial catches of marine fish and invertebrates. Canadian Special Publication of Fisheries and Aquatic Sciences, 66. 290pp

English, S., Wilkinson, C., Baker, V. (1994). Survey Manual for Tropical Marine Resources. ASEAN-Australia Marine Science Project, Australian Institute of Marine Science, Townsville, Queensland

Fabens, A. (1965). Properties and fitting of Von Bertalanffy Growth Curve. Growth 29:265-289

Fairweather, P.G. (1991). Statistical power and design requirements for environmental monitoring. Aust. J. Mar. Freshwater Res. 42:555-567

Faith, D.P., Humphrey, C.L., Dostine, P.L. (1991). Statistical power and BACI designs in biological monitoring; comparative evaluation of measures of community dissimilarity based on benthic macro-invertebrate communities on Rockhole Mine Creek, Northern Territory, Australia. Aust. J. Mar. Freshwater Res. 42:589-602

Faith, D.P., Dostine, P.L., Humphrey, C.L. (1995). Detection of mining impacts on aquatic macro-invertebrate communities: results of the disturbance experiment and the design of a multivariate BACIP monitoring programme at Coronation Hill, Northern Territory. Australian Journal of Ecology 20:167-180

Field, J.C., Clarke, K.R., Warwick, R.M. (1982). A practical strategy for analysing multispecies distribution patterns. Mar. Ecol. Prog. Ser. 8:37-52

Fishbase (1995). Fishbase: a biological database on fish. Ver 1.2 CD-ROM, Froese, R., Pauly, D., eds. ICLARM, Manila

Fowler, A.J. (1987). The development of sampling strategies for population studies of coral reef fishes. A case study. Coral Reefs 6:49-58

Francis, R. I. C. C., Paul, L. J., Mulligan, K. P. (1992). Ageing of adult snapper (Pagrus auratus) from otolith annual ring counts: validation by tagging and oxytetracycline injection. Australian Journal of Marine and Freshwater Research 43, 1069-89

Gabriel, W.L., Sissenwine, M.P., Overholtz, W.J. (1989). Analysis of spawning stock biomass per recruit: an example for Georges Bank haddock. N. Am. J. Fish. Man. 9:383-391

Garcia, S., Sparre, P., Csirke, J. (1989). Estimating surplus production and maximum sustainable yield from biomass data when catch and effort time series are not available. Fisheries Research 8:13-23

GBRMPA (1979). Workshop on coral trout assessment techniques. Held at Heron Island 21 April - 4 May 1979. GBRMPA Workshop Series No.3

Goodyear, C.P. (1989). Spawning stock biomass per recruit: the biological basis for a fisheries management tool. ICCAT Working Document SCRS/89/82. 10pp

Green, R.H. (1979). Sampling Design and Statistical Methods for Environmental Biologists. Wiley Interscience, New York

Gulland, J.A., ed. (1971). The fish resources of the ocean. (Revised edition of FAO Fish. Tech. Pap. 97). West Byfleet, Surrey, Fishing News (Books), Ltd., for FAO, Rome

Gulland, J.A., ed. (1983). Fish Stock Assessment. A Manual of Basic Methods, vol. 1. FAO/Wiley Series on Food and Agriculture, John Wiley & Sons, Chichester

Haddon, M., Willis, T.J. (1995). Morphometric and meristic comparison of orange roughy (Hoplostethus atlanticus: Trachichthyidae) from the Puysegar Bank and Lord Howe Rise, New Zealand, and its implications for stock structure. Mar. Biol. 123(1):19-27

Harmelin-Vivien, M.L., Harmelin, J.G., Chauvet, C., Duval, C., Galzin, R., Lejeune, P., Barnabe, G., Blanc, F., Ghevalier, R., Duclerc, J., Lasserre, G. (1985). Evaluation visuelle des peuplements et populations de poissons: methodes et problemes. Review of Ecology (Terre Vie) 40:468-539

Hilborn, R., Walters, C.J. (1992). Quantitative fisheries stock assessment: choice, dynamics and uncertainty. Chapman and Hall, London

Hurlbert, S.H. (1984). Pseudoreplication and the design of ecological field experiments. Ecol. Monog. 54:187-211

Jennings, S., Lock, J. (1996). Population and ecosystem effects of fishing. pp. 193-218. In: Polunin, N.V.C., Roberts, C.M., (eds.) Reef Fisheries, Chapman and Hall, London

References

Jennings, S., Polunin, N.V.C. (1995). Comparative size and composition of yield from six Fijian reef fisheries. J. Fish Biol. 46:28-46

Jennings, S., Polunin, N.V.C. (1996). Effects of fishing effort and catch rate upon the structure and biomass of Fijian reef fish communities. J. Appl. Ecol. 33:400-412

Johnson, C.R., Field, C.A. (1993). Using fixed-effects model multivariate analysis of variance in marine biology and ecology. Oceanogr. Mar. Biol. Ann. Rev. X:177-221

Kulbicki, M. (1988). Correlation between catch data from bottom longlines and fish censuses in the SW lagoon of New Caledonia. Proc. 6th Int. Coral Reef Symp. 2:305-312

Kulbicki, M., Mou Tham, G., Thollet, P., Wantiez, L. (1993). Length-weight relationships of fish from the lagoon of New Caledonia. Naga 92(2-3):26-29

Laane, W.E.M., Peters, J.S. (1993). Ecological objectives for management purposes: applying the Amoeba Approach. Journal of Aquatic Ecosystem Health 2:277-286

Lincoln Smith, M.P. (1988). Effects of observer swimming speed on sample counts of temperate rocky reef fish assemblages. Mar. Ecol. Prog. Ser. 43:223-231

Lincoln Smith, M.P. (1989). Improving multispecies rocky reef fish census by counting different groups of species using different procedures. Environ. Biol. Fishes 26:29-37

Lincoln Smith, M.P. (1991). Environmental impact assessment: the roles of predicting and monitoring the extent of impact. Aust. J. Mar. Freshwater Res. 42:603-614

Lincoln Smith, M.P., Bell, J.D. (1996). Testing the use of marine protected areas to restore and manage tropical multispecies invertebrate fisheries at the Arnavon Islands, Solomon Islands: abundance and size frequency distributions of invertebrates, and the nature of habitats, prior to declaration of the marine conservation area. Report prepared for Australian Centre for International Agricultural Research (ACIAR), Canberra, ACT, Australia

Loubens, G. (1980). Biologie quelques especes de Poissons du lagon neo-caledonian. III. Croissance. Cah. Indo-pacifique 2:101-153

Manly, B.F.J. (1991). Randomisation and Monte Carlo methods in biology. Chapman and Hall, London

Mapstone, B.D. (1995). Scalable decision rules for environmental impact studies: effect size, Type I and Type II errors. In: Detection of Ecological Impacts: Conceptual Issues and Application in Coastal Marine Habitats, pp. 67-80, Schmidt, R.J., Osenberg, C.W., eds. Academic Press, San Diego

Mapstone, B.D., Ayling, A.M. (1993). An investigation of optimum methods and unit sizes for the visual estimation of abundances of some coral reef organisms. A Report to the Great Barrier Reef Marine Park Authority. 71pp

Mapstone, B.D., Campbell, R.A., Smith, A.D.M. (1996). Design of experimental investigations of the effects of line and spear fishing on the Great Barrier Reef. CRC Reef Research Technical Report No. 7. Townsville, CRC Reef Research Centre. 86pp

Marshall, N. (1980). Fishery yields of coral reefs and adjacent shallow water environments. pp. 103-109. In: Saila, S.B., Roedel, P.M., (eds.) Stock assessment for tropical small-scale fisheries, University Rhode Island, Kingston, R.I.

McManus, J.W., Nanola, C.L., Reyes, R.B., Kesner, K.N. (1992). Resource Ecology of the Bolinao Coral Reef System. (ICLARM Stud. REV. 22), ICLARM, Manila, Philippines, 117pp

Munro, J.L., McB. Williams, D. (1985). Assessment and management of coral reef fisheries: biological, environmental and socio-economic aspects. Proceedings 5th International Coral Reef Congress, Tahiti 4:545-578

Panella, G. (1971). Fish otoliths: daily growth layers and periodical patterns. Science 137(4002): 1124-1127

Pauly, D. (1980). A selection of simple methods for the assessment of tropical fish stocks. FAO Fish. Circ., 279. 54pp

Pauly, D. (1984). Fish population dynamics in tropical waters: a manual for use with programmable calculators. ICLARM Studies and Reviews 8, Manila, Philippines

Pauly, D., Gashutz, G. (1979). A simple method for fitting oscillating length growth data, with a program for pocket calculators. ICES C.M. 1979/G:24-26

Peterman, R. (1990). Statistical power analysis can improve fisheries research and management. Can. J. Fish. Aq. Sc. 47:2-15

Plan Development Team (1990). The potential of marine fishery reserves for reef fish management in the U.S. southern Atlantic. (ed. Bohnsack, J.A.). NOAA Technical Memorandum NMFS-SEFC-261

Polunin, N.V.C., Roberts, C.M., eds. (1996). Reef Fisheries. Chapman and Hall, London

Polunin, N.V.C., Roberts, C.M., Pauly, D. (1996). Developments in tropical reef fisheries science and management. pp. 361-378. In: Polunin, J.V.C., Roberts, C.M., (eds.) Reef Fisheries, Chapman and Hall, London

Rawlinson, N.J.F., Milton, D.A., Blaber, S.J.M., Sesewa, A., Sharma, S.P. (1995). A Survey of the Subsistence and Artisanal Fisheries in Rural Areas of Viti Levu, Fiji, ACIAR Monograph No.35, 135pp. Australian Centre for International Agricultural Research, Canberra

Recksiek, C.W., Appeldoorn, R.S., Turingan, R.G. (1991). Studies of fish traps as stock assessment devices on a shallow reef in south-western Puerto Rico. Fisheries Research 10:177-197

Ricker, W.E. (1975). Computation and interpretation of biological statistics of fish populations. Bulletin of the Fisheries Research Board of Canada, 191. 382pp

Roberts, C.M., Polunin, N.V.C. (1993). Marine reserves: simple solutions to managing complex fisheries? Ambio 22(6):363-368.

Robson, D.S. (1966). Estimation of relative fishing power of individual ships. ICNAF Research Bulletin. 3:5-14

Russ, G.R. (1984). A review of coral reef fisheries. UNESCO Rep. Mar. Sci. 27:74-92

Russ, G.R. (1985). Effects of protective management on coral reef fishes in the central Philippines. Proceedings 5th International Coral Reef Symposium, Tahiti 4:219-224

Russ, G. R. (1991). Coral reef fisheries: effects and yields, pp. 601-635. In: Sale, P.F. (ed.) The Ecology of Fishes on Coral Reefs. Academic Press, Inc., San Diego

Sale, P.F., ed. (1991). The Ecology of Fishes on Coral Reefs. Academic Press, San Diego

Sale, P.F., Sharp, B.J. (1983). Correction of bias in visual transect censuses of coral reef fishes. Coral Reefs 2:37-42

Samoilys, M.A. (1992). Review of the underwater visual census method developed by the DPI/ACIAR project: visual assessment of reef fish stocks. Conference and Workshop Series QC92006. Department of Primary Industries, Brisbane, 55pp

Samoilys, M.A. (1997). Movement in a large predatory fish: coral trout *Plectropomus leopardus* (Pisces: Serranidae) at Heron Island, Australia. Coral Reefs: 16: 151-158

Samoilys, M.A., Carlos, G. (1992). Development of an underwater visual census method for assessing shallow water reef fish stocks in the south west Pacific. ACIAR Project PN8545 Final Report, April 1992. 100pp

Samoilys, M.A., Carlos, G. (in prep.) Determining optimal methods of underwater visual census for estimating the abundance of coral reef fishes.

Samoilys, M., Fuentes, H., Tuwai, I., Tikomainiusiladi, B., Leqata, J., Oreihaka, E., Mobiha, A., Potuku, T., Die, D., Connell, S., Lincoln Smith, M., Wilson, M., Watson, R. (1995). Application of underwater visual census to assessing coral reef fish stocks in the tropical pacific. Report prepared for the Australian Centre for International Agricultural Research (ACIAR). ACIAR Project Number 9304, Final Report, September 1995

Saville, A. ed (1977). Survey Methods of appraising fisheries resources. FAO Fisheries Technical Paper, 171. 76pp

Schnute, J. (1981). A versatile growth model with statistically stable parameters. Can. J. Fish. Aq. Sc. 38:1128-1140

Siegel, S., Castellan, N.J.J. (1988). Nonparametric Statistics for Behavioral Sciences. McGraw-Hill Book Co., Inc., New York

Snedecor, G.W., Cochran, W.G. (1989). Statistical Methods, 8th edn. Iowa University Press, Ames, Iowa

Sokal, R.R., Rohlf, F.J. (1981). Biometry, 2nd edn. W. H. Freeman, San Francisco

Somers, I. (1988). On a seasonally oscillating growth function. Fishbyte 6(1):8-11

Sparre, P., Venema, S.C. (1992). Introduction to Tropical Fish Stock Assessment, Part 1. Manual. FAO Fisheries Technical Paper 306/1 Rev.1. FAO, Rome

Stapleton, P. (1987). Writing Research Papers. An easy guide for non-native English speakers. ACIAR Report, Canberra

Stewart-Oaten, A., Murdoch, W.W., Parker, K.R. (1986). Environmental impact assessment: 'pseudoreplication' in time? Ecology 67:929-940

Thresher, R.E., Gunn, J.S. (1986). Comparative analysis of visual census techniques for highly mobile, reef-associated piscivores (Carangidae). Environ. Biol. Fishes 17(2):93-116

Underwood, A.J. (1981). Techniques of analysis of variance in experimental marine biology and ecology. Oceanogr. Mar. Biol. Ann. Rev. 19:513-605

Underwood, A.J. (1990). Experiments in ecology and management: their logics, functions and interpretation. Australian Journal of Ecology 15:365-389

Underwood, A.J. (1993). The mechanics of spatially replicated sampling programmes to detect environmental impacts in a variable world. Australian Journal of Ecology 18:99-118

Underwood, A.J. (1995). Ecological research and (and research into) environmental management. Ecological Applications 5:232-247

Underwood, A.J. (1997) Experiments in ecology: their logical design and interpretation using Analysis of Variance. Cambridge University Press

UNESCO (1984). Comparing coral reef survey methods. Report of a regional UNESCO/UNEP workshop, Phuket, Marine Biology Centre, Thailand. 13-17 December 1982. UNESCO Reports in Marine Science 21

Walters, C.J. (1986). Adaptive Management of Renewable Resources. MacMillan, New York

Watson, M., Ormund, R.F.G. (1994). Effect of an artisanal fishery on the fish and urchin populations of a Kenyan coral reef. Mar. Ecol. Prog. Ser. 109:115-129

Watson, R.A., Carlos, G. M., Samoilys, M.A. (1995). Bias introduced by the non-random movement of fish in visual transect surveys. Ecol. Modelling 77: 205-214

Winer, B.J., Brown, D.R., Michels, K.M. (1991). Statistical Principles in Experimental Design, 3rd edn. McGraw-Hill, New York

Wright, A., Richards, A.H. (1985). A multispecies fishery associated with coral reefs in the Tigak Islands, Papua New Guinea. Asian Mar. Biol. 2:69-84

Zar, J.H. (1984). Biostatistical analysis, 2nd edn. Prentice-Hall Inc, Englewood Cliffs, New Jersey

Zar, J.H. (1996). Biostatistical analysis, 3rd edn. Prentice-Hall International Inc, Upper Saddle River, New Jersey

Field trip equipment checklist

UVC surveys

Species identification book(s)

Datasheets printed on waterproof paper

Datasheet clip-boards and clips

Pencils and string

Stopwatch

50m vinyl tape measure

5m x 3mm floating rope with fishing float

SCUBA gear

Dinghy gear

Fishery surveys

Species identification book(s)

Datasheets pre-printed

Datasheet clip-boards and clips

Pencils and string

Measuring board (or tape)

Weighing scales

Notes

Notes

Notes

Appendix

Notes

Appendix

MANUAL
FOR
ASSESSING
FISH STOCKS
ON
PACIFIC
CORAL
REEFS

edited by

Melita Samoilys

The Department of Primary Industries has taken all reasonable steps to ensure that the information contained in this publication is accurate at the time of publication. Readers should ensure that they make appropriate enquiries to determine whether new information is available on the particular subject matter.

FOREWORD

For many Pacific Island countries fish remains the major component of the protein diet, especially in atoll and small island communities where food gathering and recreation centre predominantly on the sea and its bounty. Steadily increasing populations and a continuing trend to urban concentration have encouraged the expansion of artisanal and commercial fishing operations which in turn have contributed to a depletion of accessible finfish and other marine food resources particularly around major population centres. The growing pressure on inshore reef resources and the impact on traditional subsistence catches are matters of growing concern to many island governments, and the cause of frequent conflict between fishermen and resources owners. Recent reviews of fishery research requirements in Pacific Island countries have consistently stressed the need for priority attention to the development of national resource data bases adequate to support sound management decisions by policy makers.

Responding to requests from several Pacific Island countries for assistance in this area, and with funding from the Australian Centre for International Agricultural Research (ACIAR), a collaborative research team lead by Melita Samoilys from the Queensland Department of Primary Industries (DPI) and comprising Australian scientists and colleagues from Solomon Islands, Fiji and Papua New Guinea, worked on the development of simple methods for assessing stock status and monitoring the impacts of fishing and other factors on exploited coral reef resources. This productive study spanning several years centred on the application of underwater visual census techniques (UVC) chosen for their utility in shallow tropical coral reef environments and relevance in the Pacific Island context.

It was anticipated that the methods and approaches refined during this study would have broader application to other countries in the Pacific and elsewhere. Accordingly, one of the defined tasks for the ACIAR project team was the preparation of a practical guide which distilled the lessons and outcomes from this and past research efforts by DPI and others, to permit wide dissemination of the research results to scientists in other countries. This manual has met the challenge exceptionally well, and its contents reflect the dedicated efforts of a great many people, several of whom are recognised as key contributors to individual chapters. It is particularly pleasing to see several Pacific Island researchers listed as authors which underlines the truly collaborative effort that went into its preparation. To ensure that the contents met the requisite criteria of scientific validity, simplicity and utility, the manual was subjected to a rigorous review process. The manual usefully integrates the approaches developed during the ACIAR project with resource assessment methods emanating from earlier work in Australia and elsewhere, most notably the ASEAN-Australian Living Coastal Resources Project. The end result is a clear and practical guide to the complex task of assessing food fish stocks on coral reefs. It is appropriate that the final chapter is dedicated to the reporting of results covering both the preparation of scientific papers and the presentation of research outcomes to resource managers and to the community at large, the latter an area too often neglected in scientific endeavours.

I wish to thank the Editor, Melita Samoilys, and the individual authors, for the hard work that went into preparing the manual and for the high standards they set and maintained for its contents. I would also like to acknowledge the many scientists both in Australia and overseas who contributed in many different ways to this outstanding publication.

Barney Smith
Research Program Coordinator, ACIAR Fisheries Program

EDITOR'S ACKNOWLEDGEMENTS

CONTENTS

CONTRIBUTORS

Gary Carlos
27 Alfrison Drive
Buderim, Queensland 4556, Australia

David Die
Marine Research Division, Commonwealth Scientific
and Industrial Research Organisation,
PO Box 120, Cleveland, Queensland 4163, Australia
david.die@marine.csiro.au

Neil Gribble
Northern Fisheries Centre, Department of Primary
Industries, PO Box 5396, Cairns, Queensland 4870,
Australia
gribbln@dpi.qld.gov.au

Robert Koelldorfer
Australian Maritime College, PO Box 21,
Baconsfield, Tasmania 7270, Australia

Marcus Lincoln Smith
The Ecology Lab, 14/28-34 Roseberry Street,
Balgowlah, New South Wales 2093, Australia
ecolab@ozemail.com.au

Melita Samoilys
Fisheries, Department of Primary Industries,
PO Box 1085, Townsville, Queensland 4810,
Australia
samoilm@dpi.qld.gov.au

Krishna Swamy
Fisheries Division, Ministry of Agriculture
Fisheries and Forest,
PO Box 358, Suva, Fiji

Iliavi Tuwai
Fisheries Division, Ministry of Agriculture
Fisheries and Forest,
PO Box 358, Suva, Fiji