INTRODUCTION

Relationships between human activity and the non-human (='natural') environments have become increasingly important in recent decades (Hendee *et al.* 1990). In particular, human impacts on natural environments are seen as undesirable when those impacts deliberately or inadvertently drive natural phenomena beyond the limits expected in the absence of human intervention. Since natural phenomena cannot, in general, be managed directly, 'environmental management' (to reduce the deleterious effects of human activities) hinges on the regulation of human activities (Kenchington 1990). Implicit in such a strategy are the assumptions that: i) a managed activity does, or would in the absence of regulation, push the natural environment beyond its 'normal' behaviour; ii) the natural environment will take care of itself if human perturbations are minimised; and iii) the regulation of human activities successfully ameliorates their environmental impacts. Accordingly, assessing the success or failure of management strategies requires knowledge of: i) the normal status or behaviour of the natural environment; ii) the degree to which anthropogenic impacts force the environment beyond normal conditions; and iii) the effectiveness of management in reducing impacts. Sound information about the status and behaviour of the environment, both in the presence and absence of human activity, therefore, is essential for assessing the efficacy of management strategies (*e.g.*, see Alcala 1988, Russ 1984a, 1989). McNeill (1994) has emphasised, however, that little has been done toward monitoring the status of Marine Protected Areas in Australia, or toward assessing the effectiveness of their management.

A General Monitoring Protocol for the GBR

The gazetting of the Great Barrier Reef as a multi-use marine park explicitly demanded the conservation of the biological characteristics of the Great Barrier Reef in the context of ongoing recreational use and commercial development (GBR Marine Park Act 1975, Kenchington 1990). To ensure that all provisions of the Act are met, the Great Barrier Reef Marine Park Authority (GBRMPA) must regulate human activities to minimise impacts on the (natural) GBR environment. The favoured regulatory strategy to date has been to zone the GBR for differential access and use (Kenchington 1990, GBRMPA 1983, 1985, 1987, 1988, 1992).

These responsibilities, and the concerns of users of the reef, are manifest at a variety of scales of space and time. Assessment of specific impacts and issues of reef use are typically addressed at relatively local scales (within reefs) and over short times (one to five years). Zoning of the GBR and general management strategies, however, extend to very large spatial scales (reefs, regions) and are operative over long times (5 years - decades). Adequate judgement of management strategies with respect to conservation of the GBR environment requires sound empirical knowledge of spatial and temporal patterns in the distribution and abundance of organisms on the GBR under 'normal' conditions, the variation inherent in those patterns, and of the resilience of populations to perturbation. This information is most efficiently provided by carefully planned quantitative descriptive studies over a range of spatial and temporal scales - *i.e.*, *via* a sound monitoring programme - combined with manipulative experimental studies.

If longer term monitoring studies and local impact assessment studies are to be designed for maximum benefit at minimum cost, reliable estimates of natural variability in abundances at a range of spatial and temporal scales are needed. Armed with knowledge of natural variability in abundances, we can predict the sensitivity of monitoring programmes and their power to detect non-natural perturbations such as anthropogenic impact and the influence of various management strategies (such as zoning plans). These predictions can, and should, be tested as opportunities arise, and revised as methodology and experience improves. It is essential that the limitations of a monitoring programme (in terms of the precision of estimates and the magnitudes of differences detectable) be clearly identified so that monitoring programmes can be designed to cater for

particular objectives, and the results of those programmes can be interpreted realistically (Andrew & Mapstone 1987, Green 1979, Keough & Mapstone 1995, Mapstone 1995, 1996).

The development of a monitoring programme is most sensibly approached in three stages, neither one of which alone provides sufficient information for the adequate definition of an optimum monitoring programme. In the first stage, the relationships between methodology and small-scale biological features should be thoroughly examined, resulting in the choice of the optimum sampling unit and method of survey for each subject species or group of organisms (Andrew & Mapstone 1987, Downing 1979, Downing & Anderson 1985, Downing & Cyr 1985, Downing *et al.* 1987, Fowler 1987, Green 1979, Kenelly & Underwood 1984, 1985, Mapstone 1988, Mapstone & Ayling 1993, Pringle 1984, Sale & Sharp 1983). It should be verified that the chosen sampling unit has adequate sampling characteristics over the range of environmental conditions (*e.g.* habitat, population density) within which it will be used (Mapstone 1988, Mapstone & Ayling 1993, Lincoln Smith 1988, 1989). These aspects of sampling a number of organisms relevant to the GBR have been examined previously (Bell *et al.* 1985, Bohnsack & Banerot 1983, Brock 1982, Fowler 1987, GBRMPA 1978, 1979, 1986, Harmelin-Vivien *et al.* 1985, Kimmel 1985, Mapstone 1988, Mapstone & Ayling, 1993, Sale & Douglas 1981, Samoilys & Carlos 1992, Sanderson & Salonsky 1980, Sale & Sharp 1983).

In the second stage, the most cost-effective, least biased, and most stable sampling method is used to estimate the variation in abundances of organisms over a range of spatial and temporal scales (Caffey 1985, Doherty 1987, Eckert 1984, Keough & Mapstone 1995, Sale *et al.* 1984, Underwood 1991). Results of this stage provide the information necessary to optimise the allocation of effort to various levels in a monitoring programme such that the data obtained will provide adequate resolution and be most sensitive to changes over both time and space. The choice of scales to be considered inevitably will be arbitrary, to some extent, and/or determined by the perceived purposes of a monitoring programme, but existing knowledge of the biology of the subject organisms should also be taken into consideration (Resh 1979). In a third stage of research, the predicted performance of a suggested monitoring programme should be tested by manipulative field studies.

Random Variances & Sampling Designs

The design of a sampling, monitoring, or experimental study typically is a trade-off between desired rigour, statistical power of hypothesis tests, or precision of estimates, and the costs of doing the research (Andrew & Mapstone 1987, Peterson 1993, Warwick 1993). Refinement of the trade-off can be considered in three main steps: i) identification of the effective experimental unit at which nominated 'treatment' or systematic effects should be replicated and the most cost-effective methods for measuring effects; ii) consideration of potential sub-sampling requirements within replicate units such that the scale of the experimental units is adequately covered with the sampling method(s) given logistic and cost constraints; and iii) estimation of the numbers of experimental units that should be sampled to detect effects that are considered important with a nominated certainty. Each of these steps depends on (usually prior) estimation of variances in measured variables (*e.g.*, abundance of organisms) and the explicit consideration of the costs of sampling (Andrew & Mapstone 1987, Bros & Cowell 1987, Cochran 1963, Cohen 1988, Green 1979, Millard & Lettenmaier 1986, Underwood 1981, Winer *et al.* 1991). Ideally, pilot studies preceding each project should provide a trial ground for sampling methods and robust estimates of the costs of sampling and variances of estimates. In most situations, however, pilot studies are either small in scope or non-existent.

The appropriate experimental unit will be case specific and a matter of definition in the context of the question being asked (Andrew & Mapstone 1987, Hurlbert 1984). The choice of sampling methods should revolve around the sampling properties and logistic considerations of alternative available methods, and will impinge directly on comparisons among studies. Hence, in many instances, similar methods will be adopted in several studies. This tendency often reflects a belief

that standardisation of methods provides insurance against case-specific biases that impinge on comparability of results, rather than independent examinations of the properties of chosen methods (Andrew & Mapstone 1987). Visual surveys by divers are a popular method of quantifying abundances of demersal macro-biota in shallow reefs, and the sampling properties of several manifestations of visual survey methods have been examined in detail previously (see Andrew & Mapstone 1987 for review, Fowler 1987, Mapstone & Ayling 1993, Samoilys & Carlos 1992, Thresher & Gunn 1986).

The necessity for sub-sampling within experimental units also will generally be case specific. Choice of sub-sampling schemes will be a product of: (i) the size of sampling units relative to experimental unit; (ii) the logistic capacity to randomly distribute sampling units over the experimental units; and (iii) prior knowledge of the scales at which variation within experimental units is likely to be non-trivial, and, therefore, should be targeted specifically in order to minimise the potential for inflated variation among replicate experimental units (Cochran 1963, Cochran & Cox 1957, Keough & Mapstone 1995, Underwood 1981). Combined with known costs of sampling, variance estimates at each sub-sampling stratum can be used to predict the allocation of available resources (effort, money) among different levels in hierarchical sampling schemes such that the overall variance is minimised for a given total expenditure (Andrew & Mapstone 1987, Cochran 1963, Snedecor & Cochran 1980, Underwood 1981).

Neither the scale-related variations in abundances of demersal reef biota nor the cost-benefit relations of sampling at different nested scales within reefs have been examined widely in tropical systems (but see Doherty 1987, 1991, Fowler 1987, Mapstone 1988). Justifiable generalisations about the scales at which biota vary most or least within coral reefs will provide clear guidance for the design of future monitoring or experimental field studies, especially where extensive dedicated pilot studies are impossible. For such generalisations to be useful, however, the uncertainty in variance estimates or in 'optimum' allocations of effort to different sub-sampling strata within experimental units must be examined. This has not been done empirically in any marine systems, with the result that point estimates of variance components or sub-sampling schemes are accepted with unknown confidence.

Finally, there is increasing concern about the adequacy of replication of experimental units in ecological studies to detect effects of experimental treatments or natural phenomena that might be considered important. Several authors have recommended the consideration of statistical power when planning studies, and using power calculations to predict the amount of replication necessary to detect nominated effects (Andrew & Mapstone 1987, Bernstein & Zalinski 1983, Green 1989, Keough & Mapstone 1995, Mapstone 1995, 1996, Millard 1987, Millard & Lettenmaier 1986, Peterman 1990, Toft & Shea 1983, Underwood 1981, 1991, 1993, 1996). Again, this approach is relatively uncommon in tropical reef studies (but see Brodie *et al.* 1989, 1992, Kaly *et al.* 1993a,b, Mapstone *et al.* 1989, 1992, Mapstone 1992, Mapstone *et al.* 1994). There is potentially considerable advantage to prior derivation of estimates of the relationship between replication and detectable effects at scales that are likely to be important for future studies of, for example, management regimes or human impacts (Bence *et al.* 1996, Carney 1996, Faith *et al.* 1995, Hunphrey *et al.* 1995, Keough & Black 1996, Keough & Mapstone 1995, Osenberg *et al.* 1996, Resh *et al.* 1995, Schmitt & Osenberg 1996, Stewart-Oaten 1996, Thrush *et al.* 1996, Underwood 1993, 1996). As with cost-benefit analyses, however, the uncertainty in predictions of required replication is rarely considered.

In this study we investigated variability in the abundances of a number of reef organisms at a range of spatial scales in the interests of seeking some general empirical bases for the design of future sampling and monitoring studies. We examined estimates of variances at a hierarchy of spatial scales known to be of interest for a variety of coral reef studies, including fundamental research, management strategy evaluation, and assessments of environmental impacts. We used cost-benefit analyses to consider empirically the potential for generalisation in suggested allocations of effort to sub-sampling at different spatial scales, and the precision of those estimates given the sort of pilot data that would be available in most studies. Finally, we used our estimates of variances to

predict the replication necessary to detect hypothetical effects on reef biota at three scales, and consider empirically the uncertainty in those predictions.

We were concerned principally with:

- *Acanthaster planci*, *Linckia laevigata*, and *Tridacna* spp.;
- Sessile benthic biota and non-living substrata, with particular emphasis on live corals;
- Fish with medium to great mobility over short periods, including *Plectropomus* spp., lutjanids, chaetodontids, and lethrinids;
- Fish with restricted home-ranges and relatively low mobility over short intervals, such as most of the pomacentrids and some labrids.

We chose to cover as many organisms as logistically possible because: 1) a general monitoring programme should take into account the status of several species; 2) the optimum sizes of sampling units proved to be the same for several organisms (Mapstone & Ayling 1993); 3) many of the organisms can be efficiently counted concurrently; and 4) much of the cost of such a study is incurred in getting to survey sites and support costs whilst in the field, and it was therefore desirable to maximise the return from such costs.