International<br>Whaling<br>Commission

# The Comprehensive Assessment of Whale Stocks: the early years 

## Edited by

G. P. Donovan

The International Whaling Commission,
The Red House, Station Road
Histon, Cambridge CB4 4NP
United Kingdom

## Reports of the International Whaling Commission Special Issue 11

Cambridge 1989

# The International Whaling Commission <br> was constituted under the International Convention for the Regulation of Whaling signed at Washington on 2 December 1946. 

ISSN 0255-2760
ISBN 0906975220

The International Whaling Commission
The Red House, Station Road, Histon,
Cambridge CB4 4NP

## Preface

At its 1982 meeting in Brighton, the International Whaling Commission adopted a proposal to include in the Schedule to the Convention ${ }^{1}$ a pause in commercial whaling (commonly known as the 'moratorium') to take effect from the 1985/86 pelagic and 1986 coastal seasons. The initial proposal simply stated that 'catch limits for the killing of whales for commercial purposes shall be zero' but before the vote was taken some extra wording was added, perhaps to indicate to the whaling countries that the proposal seriously considered the possibility of whaling resuming.

## The extra wording was as follows:

'This provision will be kept under review, based upon the best scientific advice, and by 1990 at the latest, the Commission will undertake a comprehensive assessment [my italics] of the effects of this decision on whale stocks and consider modification of this provision and the establishment of other catch limits'.
However, there was no discussion at all at that meeting as to the meaning of 'comprehensive assessment' even though the wording was adopted and became part of the Commission's Schedule to the Convention - essentially part of a legal document.

At the next Commission Meeting, as part of a Japanese initiative, the Commission agreed with a Scientific Committee proposal that it should begin to plan for a still undefined comprehensive assessment by looking at problems in current information on whale stocks and what 'conceptual' approaches might be used to provide the Commission with more effective advice ${ }^{2}$.

At the same Meeting, the Commission added an extra piece to the puzzle by introducing a new management scheme for aboriginal subsistence whaling which also included the words 'comprehensive assessment', this time stating that the scheme would be

> 'kept under review ... and by 1990 at the latest, the Commission will undertake a comprehensive assessment of the effects of these provisions on whale stocks and consider modification'.

In 1984 the Scientific Committee again indicated that it did not understand what the Commission intended to comprise a comprehensive assessment. However, despite this, it discussed what it thought the Commission might mean, particularly in the context of problems it was facing in trying to give the Commission advice. Very briefly these centred on certain management problems including acceptable risk levels in the context of maximising yield and preventing-over exploitation; methodological concerns about the ways being used to estimate abundance and biological parameter values; and the relationships between management policies and assessment methods.

Despite a further request that the Commission should explain what was meant by a comprehensive assessment, at the 1985 Scientific Committee meeting the Committee was still in the position of having no Commission guidance. It therefore decided that if progress was to be made it would have to define what it thought was a 'comprehensive assessment' and establish how it might be accomplished.

[^0]To this end it recommended, and the Commission agreed, to hold a special meeting on the subject.

That meeting (the report of which is reprinted in this volume) was held in April 1986 and it was agreed that from a Scientific Committee viewpoint, the Comprehensive Assessment ${ }^{3}$ can be considered as an in-depth evaluation of the status of all whale stocks in the light of management objectives and procedures and that this would include the examination of current stock size, recent population trends, carrying capacity and productivity. Three major areas of work were identified:
(1) to review and revise current knowledge concerning methodology, stock identity and data availability;
(2) to plan and conduct the collection of new data;
(3) to examine alternative management regimes.

These three areas of course are heavily inter-related and particularly important is the relationship between management regimes and data and methodology requirements. The Scientific Committee has identified major problems with the Commission's current management procedure, which is tied to the concept of MSY and a need to have estimates of initial and current population size as well as a series of other parameters in order to determine population trajectories and MSY levels. An integral part of the Comprehensive Assessment was the encouragement and funding of a series of simulation studies of alternative management procedures. Five papers outlining progress on these studies as well as the reports of two workshops on this subject are included in this volume.

Another major area of concern is that of stock identity. Of course, if the aim is to assess the abundance and dynamics of a 'stock' then one must know what comprises that 'stock' - in geographical or other terms. It has to be said that at the moment we do not have a good idea of the biological stock identity of many of what we term 'management stocks'. The past methods used to determine 'stock' boundaries have included examination of catch distributions, adoption of fisheries boundaries, apparently 'pure' guess work and examination of the movements of marked whales. There are two approaches to this problem. One is to simulate the effects of possible boundary errors and see if they really are important to management (this is discussed in the Report of the Workshop held in Lowestoft in February 1989 and included in this volume). The other is to determine an appropriate methodology to determine biological 'stocks' and collect the relevant data. In fact what is required is a combination of both approaches. In this regard the International Whaling Commission commissioned the review of the use of new molecular techniques to examine stock identity questions that is published in this volume. This review led to further work and a Workshop was held in La Jolla in September 1989. The Report of the Workshop and associated papers will comprise a further volume in our special issue series, which is expected to be available before the end of 1990.
${ }^{3}$ Instant 'tradition' has resulted in the Comprehensive Assessment, as defined by the Scientific Committee, to be given capital initials. It is broader in scope than the comprehensive assessments referred to in the Schedule but should, if and when completed, provide the information to carry out the review required by the Schedule.

As part of its initial examination of methodology, the Scientific Committee examined the question of the estimation of current numbers (in particular, census techniques) and the use of mark-recapture data. The major review of census techniques commissioned is published in this volume. Large numbers of whales had been marked using Discovery marks (steel tags fired into a whale and recovered from the carcass) but no rigorous analyses of these data had been carried out. Initially the International Whaling Commission commissioned a review of the applicability of the whale data to current mark-recapture models and that review found that open population models and in particular Jolly-Seber type models held the most potential ${ }^{1}$. On the basis of that review, the study published here, an analysis of the best mark-recapture data available (that for the minke whales in the Southern Hemisphere), was undertaken.
The final methodological area addressed thus far is that of estimating trends in abundance. Classical fisheries theory has emphasised the use of CPUE (catch per unit effort) data and it has historically been a major element in the assessments of many whale stocks. However, surprisingly little attention had been paid, at least in the cetacean literature, to the key assumptions behind this, i.e. that CPUE is firstly a true index of abundance in the area where whaling has occurred and secondly that can this be extrapolated to the total stock area. The report of a workshop held in Reykjavik in March 1987 is reprinted in this volume. In brief it found that there really are no models which suitably mimic the relationship of CPUE and abundance. Indeed, given the huge variation in operational factors, both with time within an operation and among operations, a detailed model needs to be developed for each particular fishery if CPUE data are to be used. Put another way, apart from giving a very gross picture where there is a major crash in a population, the inherent variability in CPUE data means that, at present, they are unlikely to be useful in assessing trends in population size a conclusion which has ramifications in several other fisheries situations.

An important factor in what might be termed pre- and post-moratorium whale science is that many of the 'classical' methods of estimating abundance and biological parameters such as age at sexual maturity, mortality rates, etc. depended on information collected from dead whales. Irrespective of the value of the resultant estimates for management, the availability of samples is now clearly limited. Over the last 10 years several exciting non-lethal techniques have been developed which enable information required for management to be obtained for at least some species and populations. The Commission recognised that these techniques must play a part in the Comprehensive Assessment programme and sponsored the symposium and workshop held in La Jolla in April 1988, which concentrated on the use of such techniques with regard to individual identification of whales (primarily by photo-identification) to obtain information needed for

[^1]management. The fruits of this workshop will be revealed in a further volume in our special issue series due out in Spring/Summer 1990.

In conclusion, although initially work on the Comprehensive Assessment began slowly (almost four years elapsed before its planning got underway), I think the progress made since then has surprised many who thought it might become just another unfulfilled 'initiative'. There is still much work to do and the fact that it will not be completed by 1990 for all stocks is apparent from the 1989 Report of the Scientific Committee: by the end of the 1990 meeting, at best in-depth evaluations will only have been carried out for three groupings - the eastern North Pacific gray whales, the Southern Hemisphere minke whales and the North Atlantic minke whales. The present timetable for adopting a new management procedure does not envisage one before mid-1991. However I hope the papers in this volume emphasise the value of the work done so far and I look forward to the production of 'The Comprehensive Assessment of Whale Stocks: the Final Years'.

There are inherent difficulties in producing a volume such as this which details work carried out as part of a long-term programme. It is for this reason I have included acceptance dates for each paper. Inevitably while waiting for sufficient papers to fill a volume some of the earliest studies will be overtaken by events. This is particularly true with respect to papers describing the development of management procedures to be found in the final section of this volume. However I believe it is important to document the progress made thus far: it represents some of the most innovative work in the field of theoretical wildlife management and being able to follow the development of various procedures, including their cul-de-sacs, is almost as scientifically valuable as reviewing the final developed versions.

Similarly, although publication of the genetics techniques review by Hoelzel and Dover was held up while waiting for other papers in the volume, it remains as an excellent introduction to a previously neglected area of cetacean management biology and describes relevant developments up to mid-1988 in one of the most exciting fields of biological research today. Only with such a background can cetacean biologists not specialising in genetics appreciate the significance of the more recent developments discussed at the Workshop held in La Jolla to which I referred earlier.

Finally I would like to thank those people who helped in the production of this volume: firstly, the authors of the papers, particularly those whose work was completed in 1988, for their patience and diligence; secondly, those scientists who acted as anonymous reviewers for the papers in the volume - a thankless but vital task; and, to use the old cliché, last but not least Stella Duff and Helen Coulson, who between them typed all the text, the tables and proof-read each paper at least three times, a soul-destroying task they accomplished with their usual efficiency and cheerfulness.
G.P. Donovan

Cambridge
November 1989.

# International Whaling Commission Special Issue Series 

1 [1977]—Report of the Special Meeting of the Scientific Committee on Sei and Bryde's Whales. i-v + 150 pp.; ISBN 0906975034

2 [1980]-Sperm Whales.
i-iv +275 pp.; ISBN 0906975018
3 [1980]-Age Determination of Toothed Whales and Sirenians.
Eds W. F. Perrin and A. C. Myrick.
i-viii + 229 pp.; ISBN 0906975050
4 [1982]-Aboriginal/Subsistence Whaling (with special reference to the Alaska and Greenland Fisheries).
Ed. G. P. Donovan.
i-v + 86 pp.; ISBN 0906975093
5 [1983]-Special Issue on Historical Whaling Records.
Eds M. F. Tillman and G. P. Donovan.
i-v + 490 pp.; ISBN 0906975115
6 [1984]-Reproduction of Whales, Dolphins and Porpoises.
Eds W. F. Perrin, R. L. Brownell, Jr and D. P. DeMaster.
i-xii + 490 pp.; ISBN 0906975077
7 [1986]-Bibliography of Whale Killing Techniques.
By E. D. Mitchell, R. R. Reeves and A. Evely.
i-v + 162 pp.; ISBN 090697514 X
8 [1986]-Behaviour of Whales in Relation to Management.
Ed. G. P. Donovan.
i-v + 282 pp.; ISBN 0906975158
9 [1988]-The Biology of the Genus Cephalorhynchus.
Eds R. L. Brownell Jr and G. P. Donovan.
i-ix + 344 pp.; ISBN 0906975174
10 [1986]-Right Whales: Past and Present Status.
Eds R. L. Brownell Jr, P. B. Best and J. H. Prescott.
i-viii + 289 pp.; ISBN 0906975166

Available from the office of the Commission, The Red House, Station Road, Histon, Cambridge CB4 4NP, United Kingdom.

## Contents

Page
Reports of Meetings and Workshops ..... 1
Report of the Special Meeting of the Scientific Committee on Planning for a Comprehensive Assessment, Cambridge, 7-11 April 1986 ..... 3
Report of the Comprehensive Assessment Workshop on Catch per Unit Effort (CPUE), Reykjavik, 16-20 March 1987 ..... 15
Report of the Comprehensive Assessment Workshop on Management, Reykjavik, 23-25 March 1987 ..... 21
Report of the Comprehensive Assessment Workshop on Management Procedures, Lowestoft, UK, 6-10 February 1989 ..... 29
Reviews and Studies funded as part of the Comprehensive Assessment ..... 45
Hiby, A. R. and Hammond, P. S. - Survey techniques for estimating abundance of cetaceans ..... 47
Hoelzel, A. R. and Dover, G. A. - Molecular techniques for examining genetic variation and stock identity in cetacean species ..... 81
Buckland, S. T. and Duff, E. I. - Analysis of Southern Hemisphere minke whale mark-recovery data ..... 121
Papers describing potential Management Procedures ..... 145
Cooke, J. G. - Simulation studies of two whale stock management procedures ..... 147
De la Mare, W. K. - Further simulation studies on management procedures ..... 157
Magnusson, K. G. and Stefansson, G. - A feedback strategy to regulate catches from a whale stock. ..... 171
Punt, A. E. and Butterworth, D. S. - Results of first stage screening trials for a proposed whale stock management procedure ..... 191
Sakuramoto, K. and Tanaka, S. - A simulation study on management of whale stocks considering feedback systems. ..... 199

Reports of Meetings and Workshops

# Report of the Special Meeting of the Scientific Committee on Planning for a Comprehensive Assessment of Whale Stocks 

The meeting was held from Monday 7 April - Friday 11 April 1986 at the Royal Cambridge Hotel, Cambridge under the Chairmanship of Dr G.P. Kirkwood. A list of participants is given in Annex A.

## 1. CHAIRMAN'S WELCOME AND OPENING REMARKS

The Chairman welcomed the Committee members and invited participants to the meeting and drew attention to the terms of reference proposed by the Committee at its last meeting and agreed by the Commission that:
the Scientific Committee hold a special meeting to identify specific tasks, assign priorities and establish a timetable for undertaking a comprehensive assessment of whale stocks. More specific objectives will need to be developed, but would include:

- establishment of priorities for providing advice to the Commission;
- identification of specific reviews and other studies of existing information or assessment techniques required;
- establishment of requirements for new information for assessment, and identification of surveys or other work to be undertaken to provide that information;
- establishment of a timetable that will allow timely advice to the Commission;
- examination of the likely costs of the proposed programme;
- exploration of new management regimes.


## 2. APPOINTMENT OF RAPPORTEURS

It was agreed that Donovan would act as rapporteur with assistance from various members of the Committee where appropriate.

## 3. ADOPTION OF AGENDA

The Agenda adopted is given in Annex B.

## 4. ARRANGEMENTS FOR MEETING

The Committee agreed to a work schedule outlined by the Chairman.

## 5. REVIEW OF AVAILABLE DOCUMENTS AND REPORTS

A list of documents is given in Annex C. Reports of earlier discussions of the Committee on this topic (e.g. in Rep. int. Whal. Commn 35 and 36) and the Report of the Commission's Working Group on the Comprehensive Assessment of Whale Stocks (IWC/36/16) were also available.

## 6. DEFINITION OF A COMPREHENSIVE ASSESSMENT

The Committee noted that its terms of reference given under Item 1 were much wider in scope than the 'comprehensive assessment' of the effects of the decision by the Commission to set catch limits at zero, which was to be 'undertaken by 1990 at the latest' (Schedule Paragraph $10(\mathrm{e})$ ) and the 'comprehensive assessment' referred to in Schedule Paragraph 13(a)(3).

In addition it noted the terms of reference of the Joint Working Group on Comprehensive Assessment of Stocks which met in 1984 (IWC/36/14):
(a) to consider, in the light of the current information on whale stocks and the degrees of uncertainty that exist concerning some of the data and methods used, what conceptual approaches might be used to provide the Commission with more effective scientific advice and recommendations for management;
(b) to determine the studies required to implement these approaches; and
(c) to establish a time-table for the in-depth assessment of whale stocks which should be completed for major stocks currently exploited as soon as practicable.

### 6.1 Definition of a comprehensive assessment

Given the above, the Committee agreed that comprehensive assessment can be considered as an in-depth evaluation of the status of stocks in the light of management objectives and procedures. This could include examination of current stock size, recent population trends, carrying capacity and productivity.

In order to achieve this the Committee agreed that it would need to:
(a) review and revise assessment methods and stock identity; review data quality, availability requirements and stock identity;
(b) plan and conduct the collection of new information to facilitate and improve assessments;
(c) examine alternative management regimes.

As discussed later in the report, the Committee sees the carrying out of the Comprehensive Assessment as an iterative process, with considerable interaction between results from (a), (b) and (c).

Ivashin believed that a fuller definition should eventually be formulated and that this might best be achieved at the 1986 Scientific Committee meeting.
6.2 Relationship between management policies and advice required
The Committee has already noted the relationship between management policies and advice required - and hence assessment techniques which must be used (Rep. int. Whal. Commn 35: 36). This has been brought out in
examinations of the New Management Procedure and provision of advice in accordance with it (e.g. Rep. int. Whal. Commn 36: 37; de la Mare, 1986) and in particular problems encountered in estimating MSY, MSYL and initial stock levels. A comprehensive assessment will include an examination of management schemes in the light of the ability to estimate population sizes and parameters, taking into account the inevitable uncertainty in such estimates. This is discussed further under Item 8.

### 6.3 Information required for a comprehensive assessment

 The types of information required for the various elements of a comprehensive assessment are identified under the appropriate Agenda Items.
## 7. REVIEW OF CURRENT METHODS FOR STOCK ASSESSMENT AND PROVISION OF ADVICE ON CATCH LIMITS

### 7.1 Stock/management units

Several papers (SC/A86/CA1, CA3, CA4, CA9) noted the uncertainty over stock identity and the relationship between 'biological' and 'management' stocks. From a management viewpoint, what is important is the effect on stock size and on yields of any discrepancies between management boundaries and 'biological' boundaries. An examination of this problem should consider the following questions.

Q(i) Supposing we had perfect information, what would be an appropriate definition of a 'stock' for management purposes? The answer to this might be a single definition of discrete stocks, in the traditional manner, or it might involve a more complex definition involving stocks and substocks.

Q(ii) What information would be needed to identify the "stocks" defined in the answer to $\mathrm{Q}(\mathrm{i})$ : which of this information is practicable to obtain, and how?
$Q$ (iii) In the light of the answer to $Q$ (ii), how should the answer to $\mathrm{Q}(\mathrm{i})$ be modified to provide a working definition which is feasible to implement? This definition might require certain specific kinds of data that must be obtained to implement the definition. Alternately, it may be an adaptive definition that can be applied with minimal information, and which is modified as extra information becomes available.

Q(iv) What are the consequences of using the answer to (iii) for management purposes? Simulation studies could play a role in the analysis of this question.
$\mathrm{Q}(\mathrm{v})$ Where the analysis of $\mathrm{Q}(\mathrm{iv})$ identifies potential management problems that would be reasonably likely to arise in practice, what additional management safeguards could ameliorate them?

### 7.1.1 Biological stock boundaries

(1) Discovery-type marking data

Marking of whales using Discovery-type marks has been carried out for many years under national programmes and under the International Marking Scheme. The Committee believes that an examination of these marking and recapture data in conjunction with associated catch and effort data may provide valuable information on questions of not only location of boundaries, but also rates of mixing of whales across boundaries. Such analyses can only be carried out efficiently when the relevant data are encoded in a computer data-base. Much of the catch data has been
encoded by the Secretariat. However marking and effort data exist in several formats (e.g. original record sheets, published summaries) and in various locations (e.g. in national archives, in the International Marking Scheme records in the UK). The Committee recommends that the Secretariat determines the extent of such data, its format and whereabouts, with a view to encoding these data. It is important to ensure that any requests for information on national marking programmes made to member governments clearly specify the type of information required. For the purposes of this analysis, crude effort data (for example, gross or net catcher days worked by $10^{\circ}$ square) should be sufficient (but see below with respect to assessment techniques). The Committee noted the low level of marking carried out in lower latitudinal waters (i.e. breeding areas); the effect of lack of information on breeding stocks, particularly for the balaenopterid whales, will be examined under the questions given above.

## (2) Natural marking data

Studies using natural markings have shown that repeated and long-term identification of individual whales is possible for several species. In some cases, the data have shown movements within and between different feeding and breeding aggregations. Where extensive data sets exist, or could be collected, these could provide valuable information on stock boundaries and rates of mixing across boundaries.

## (3) Electrophoretic and related studies

Several workers have examined stock identity questions using electrophoretic and related techniques (e.g. Fujino, 1960, for sperm and fin whales; Wada, 1983 and 1984, for minke and Bryde's whales) and work is currently being undertaken in Japan and Iceland. Sigurjonsson noted that results of work on North Atlantic minke and fin whales will be available in the not-too-distant future. Any review of stock identity questions should include an examination of the interpretation of the results of such studies. It was noted that methods for taking suitable samples from living whales are being developed (e.g. Aguilar and Nadal, 1984; Payne, pers. comm.).

## (4) Morphometric studies

Some work has been carried out on cetacean morphometrics with respect to stock identity (e.g. Schnell, Douglas and Hough, 1985; Bushuev and Ivashin, 1986) and in the past such studies have been important in the discrimination of sub-species (e.g. Ichihara, 1964). The discriminatory power and interpretation of such work needs further analysis but it seems clear that systematic studies are required.

### 7.1.2 Examination of management boundaries

A proposed simulation study which would in part answer Q (iv) was prepared for the meeting (Annex D ). Depending on the outcome of this and similar studies, further work may be needed with respect to determination of management boundaries.

### 7.2 Available data

The Committee agreed that an essential part of a review process involved compiling and updating catalogues of available data by management unit and species, which
should include information on years available, format, accessibility and location. This is developed further in Item 10 and Annex E.
7.3 Assessment techniques and 7.4 Additional reviews/studies of existing data or techniques required

The Committee noted that there are several general areas of methodology which require review including the estimation of current stock size (e.g. by sightings, mark-recapture), the use of indices of abundance (e.g. CPUE) in determining population trends, the estimation of biological parameters (e.g. age at sexual maturity, pregnancy and natural mortality rates) etc. These are listed and discussed more fully under Agenda Item 10.

## 8. EXPLORATION OF MANAGEMENT REGIMES

Since the late 1970s, the Scientific Committee has noted problems in trying to fully implement the New Management Procedure (e.g. Rep. int. Whal. Commn, 29: 99-101; Ibid, 32: 47-8; Ibid 35: 36-7). Many of these problems are associated with difficulties in estimating MSY, MSYL and initial stock sizes. In recognition of these, the Scientific Committee began a series of meetings on revised management procedures. Subsequently the Commission established a Joint Working Group of the Scientific and Technical Committees which in 1981 developed the following objectives for any future management policy (IWC/33/13):
(1) to ensure that the risks of extinction to individual stocks are not seriously increased by exploitation;
(2) to maintain the status of whale stocks so as to make possible the highest continuing yield so far as the environment permits;
(a) never to move individual harvested stocks or groups of stocks of the same species in a direction which reduces its or their combined sustainable yield;
(3) to ensure the maintenance and orderly development of the whaling industry.

The importance of management procedures to the Comprehensive Assessment has already been noted (Item 6.2 ), and Fig. 1 illustrates the relationship between the development of revised management procedures and the results of the review of methodology. This is particularly relevant in terms of evaluating the importance of weaknesses in methodology and uncertainties in estimated quantities, such as population size or biological parameters. This will enable priorities to be set for future studies or data collection. Conversely, input from the reviews of methodology is vital to the development of practical management procedures.

The Committee noted that considerable advances in approaches to developing revised management procedures had been made since the series of meetings held on this topic in 1980 and 1981, particularly in the area of simulation studies. Examples of the types of new approaches have been presented in SC/A86/CA5, CA6 and CA9.

In SC/A86/CA9, Cooke presented some preliminary simulation studies of a management procedure regime involving continual adjustment of the catch based on the results of continuous monitoring of a stock. This procedure does not require an initial accurate estimate of sustainable yield. The preliminary results, using working values of the MSY exploitation rate coupled with conservative estimates of stock size, are encouraging.

In SC/A86/CA6, Tanaka investigated, using theoretical analysis and simulation techniques, the performance of a feedback procedure for determination of a catch limit that required estimates only of the present stock level in relation to the target level and the current rate of change of stock size. This method did not require knowledge of population models, replacement yields or absolute stock size. However some estimate of absolute abundance may be necessary to set the target level. Again, preliminary results indicated that the performance of this method was encouraging.

The Committee noted that although it could undertake a considerable amount of theoretical work itself, at certain


Fig. 1. Elements of the Comprehensive Assessment. $O$ procedure, $\square$ product (also in Figs 2 and 3).
stages in its exploration of revised management procedures, consultation with and guidance from the Commission would be essential. This might take the form of a joint Scientific and Technical Committee Working Group (see Item 10 , outline B). It also recognised, of course, that adoption or rejection of any proposed management procedure was ultimately the decision of the Commission.

## 9. NEW INFORMATION NEEDED FOR A COMPREHENSIVE ASSESSMENT

### 9.1 New data

The Committee noted that some of the requirements for new data could be listed almost immediately. However, a more comprehensive list, with associated priorities for collection of these data, could only be developed after the compilation of an inventory and an initial review of methodologies. Further requirements would be identified as the Comprehensive Assessment process progressed. Such data requirements may thus arise:
(1) out of a need to examine particular methodologies;
(2) in order to provide data necessary for a satisfactory assessment of a particular stock.

The Committee agreed, however, that if current trends in populations were to be examined as a high priority, it was important to recognise the need for instigating or continuing monitoring studies (e.g. sightings surveys) as soon as practicable.

Several members referred to the fact that if catches were to continue, for example under objection, it might be helpful for the Committee to offer advice to national groups concerning data requirements, although others noted that the Committee would not be in a position to do this until specific questions to be answered had been formulated during the Comprehensive Assessment process.

The potential contribution of data obtained from new techniques such as satellite-tracking and acoustic censusing was also discussed.

### 9.2 Alternative assessment techniques

Similarly to Item 9.1 above, the Committee concluded that the need for alternative assessment techniques would arise out of the review of methodology and data.

## 10. PLANNING FOR A COMPREHENSIVE ASSESSMENT

In its discussions of a comprehensive assessment the Committee identified two major areas of interest, 'data and methodology' and 'management procedures'. It also noted that the Comprehensive Assessment was an iterative process with considerable interaction between these broad areas. Figs $1-3$ show how the Comprehensive Assessment may be structured, illustrating the relationships between the various activities. They supplement the outline work plan given below. It should be noted that Fig. 1 does not distinguish between steps which apply to all species (e.g. review and revision of methodology) and those which result in an accepted assessment of a single stock.

## A. Outline Work Plan of Scientific Work to be Carried Out as Part of a Comprehensive Assessment

1. Develop an inventory of our current knowledge on the status of stocks. This involves giving general answers to the following groups of questions for each stock/regional grouping.
(a) Are there estimates of (i) current (within the last 10 years) stock size, (ii) initial stock size, (iii) replacement or sustainable yields, (iv) recent trends in stock size (over the last 10 years)?
(b) For each quantity (i) by what method were the estimates obtained, (ii) are there associated estimates of reliability?
(c) What data are available (as specified in Annex E)?
2. (a) Identify methodological problems involved in the determination of stock identity and the estimation of population size, trends, productivity and carrying capacity.

It is expected that this will require the consideration of some or all of the following topics: age determination and analysis; pregnancy, maturity and other biological parameters; use of marking and natural marks; analysis of catch per unit effort data; use of sightings and direct censuses.
(b) Determine, in the light of current management requirements, those aspects of assessment methodology for which improvements are most needed; and identify the actions required to achieve improvements.

The determination of critical improvements might require carrying out simulation studies, as well as the examination of past experience of developing assessment advice.


Fig. 2. Data.


Fig. 3. Methods.

STEPS REQUIRED
(1) Identify the work needed and the priorities for different studies (Scientific Committee).
(2) Prepare comments on implications for current management procedures (Scientific Committee).
(3) Implement workshops, contracts and other studies
3. (To be carried out in parallel with 2) Examine the availability of data, the extent to which existing data are not readily accessible, and the quality of data.

Taking account of the data needs of current and possible new assessment methodologies, determine
(a) what existing data should be made more readily accessible, e.g. by inclusion in the IWC database;
(b) additional data that should be collected from existing whaling operations, or as research work.

## STEPS REQUIRED

(1) Decide on priorities for data collection and compilation (Scientific Committee).
(2) Put existing data into more accessible form (Secretariat).
(3) Collect new data (countries).
4. Review scientific aspects of alternative management procedures in the light of the identified strengths and weaknesses of assessment methodologies and data.

## STEPS REQUIRED

(1) Obtain objectives for management from outline B below.
(2) Review management procedures.
(a) Identify potential management procedures: (i) management advice and decision procedures; (ii) advice required for reaching management recommendations.
(b) Determine whether available methods of estimating parameters can meet management objectives under given management procedures. This will be carried out (e.g. by simulation studies) in the light of information on the available levels of precision of estimates from the methodological review.
(c) If the answer to (b) is no then: (i) can parameter estimates be improved (data/methods)? (ii) can management procedure or decision rules be modified? (iii) if the answer to (i) or (ii) is yes then: re-evaluate from step (a). (Otherwise, a major task has gone as far as feasible).
(d) If the answer to (b) is yes then a major task has been completed.

## 5. Prepare 2nd round inventory.

The second round would provide more detailed and updated information on the actual figures for the current population numbers, yields (or net recruitment rates), trends, etc. Priorities for additional assessments that might be done in preparation for this might need to be set in the light of existing commitments of the Committee, and of the value of studies of particular stocks as case studies, or of their general ecological interest (e.g. Southern Ocean large baleen whales).
6. Examine the general features of whale dynamics, through the comparison of the information from different stocks, with the aim of improving the assessment and management advice for each stock.

This would probably involve the detailed study of a number of specific topics, which might include some or all of the following:- the variability of whale stocks, and their response to changes in their environment, including the influence of human factors, other than direct exploitation; the pattern of similarities and differences within and between species in their population parameters; the question of minimum stock size; density dependence; carrying capacity, and the problems of attaching confidence limits to estimates.

## STEPS REQUIRED

(1) Identify priority topics, and methods (e.g. workshops, contracts, etc) of tackling them (Scientific Committee).
(2) Prepare revised comments on implications for current, or possible new, management procedures (Scientific Committee).
(3) Implement workshops, contracts, etc.

## 7. Prepare 3 rd round inventory.

The estimates of current population numbers, trends, yields, etc, obtained in this round will provide the basis for the report on the Comprehensive Assessment. It will therefore be necessary to identify those stocks for which assessments are likely to be particularly important, and which, therefore, should receive priority attention.

## B. Outline Work Plan for Joint Scientific and Technical Committee Activity in Relation to a Comprehensive Assessment

(1) Review the management objectives of the Commission, as set out in previous reports (e.g. IWC/33/13).
(2) Examine current management procedures, the scientific and technical problems encountered in applying these, and the implications of these problems in achieving the objectives of the Commission.
(3) Review the scientific implications of alternative management procedures [see Item 4 of outline $A$ above].
(4) Consider other implications (operational, economic, legal, etc) of alternative management procedures in the light of the Commission's objectives.
(5) Prepare, as appropriate, suggestions for revisions to the Commission's management procedures.
10.1 Priority stocks and 10.2. Priorities for review of existing information and collection of new information
At each stage in the process of carrying out the Comprehensive Assessment, as indicated in the Work Plan above, there is a need to set priorities. These may be priorities for particular methodological studies (step A2), for compilation or collection of data (step A3) or for stock assessments (steps A5, A7). The Committee agreed that in most circumstances, the criteria for setting priorities should be dictated by the specific needs at that stage of the Comprehensive Assessment. In many cases, these criteria may indicate examination of stocks or species that would not have been considered 'priority stocks' during previous Committee meetings, where priority for assessment was given to stocks from which catches currently were being taken. Examples of the criteria for setting priorities at each step of the Comprehensive Assessment are given in the Work Plan. However, for steps A5 and 7, or where there were choices amongst equals (e.g. for which of a number of stocks should estimates of current stock size be obtained), it may be appropriate to use more traditional criteria, such as data availability and the likelihood of exploitation in relation to the current relative status of a stock.

### 10.3 Computing needs

The Committee noted with regret the imminent resignation of Dr C.A. Free from his position of Senior Analyst/Programmer with the IWC Secretariat. Especially in view of the additional workload implicit in the carrying out of the Comprehensive Assessment, it is essential that appropriate computing expertise continues to be available within the Secretariat.

In SC/A86/CA7, Free reported the current status of the IWC data coding project, which included coding of BIWS catch data, and marking data, principally from the International Marking Scheme. Progress on the coding of BIWS data has been much greater than had earlier been anticipated, and less than one year's work remains to complete this part of the project, while coding of all International Scheme marking data is expected to take 9
months. However the Committee noted that a considerable amount of catch data (some 600,000 animals) is not included in the BIWS data. The full extent and whereabouts of these data will become apparent after the compilation of the data inventory (Annex E).

The Secretary indicated that, as a result of the systems introduced by Free, the Secretariat currently has the expertise to cope with the existing data coding programme and even an expanded load (possibly with the hiring of temporary staff). Thus he anticipated that the Secretariat would be able to handle the expected additional data entry activity arising from the Comprehensive Assessment. However, Free's resignation will leave a gap in the technical and innovative programming capability of the Secretariat, which may need to be filled. Before doing this, however, it is necessary for a detailed job description to be developed as soon as possible. A decision to find a replacement can only be taken by the Commission at its Annual Meeting.

While recognising the urgency in determining whether a replacement was necessary and the nature of the duties required, the Committee agreed it could not take a decision at this meeting. The earliest point at which it could determine future computing needs was after step 1 of the Work Plan. The initial inventory (comprising general answers to A1(a) and (b) in the outline plan above) is scheduled to be available at the 1986 annual meeting of the Committee (see Item 10.4 below), and it is recommended that this be considered as a matter of urgency at that meeting.

### 10.4 Timetable

The Committee agreed that completion of a comprehensive assessment for every whale stock would take a considerable length of time. Consequently a timetable was developed with a view to completion of the Comprehensive Assessment for at least the major stocks, which could be included in an interim report prepared by 1990.

Because of the iterative nature of the Work Plan, with identification of priority studies being made at various stages of the process, it is very difficult to present a detailed timetable. Also, the speed with which the various items of work can be carried out depends critically on the available resources (both manpower and financial). In the Work Plan, particular tasks that could clearly be assigned either to the full Scientific Committee or the Secretariat are so designated. For the remaining tasks the Committee identified a number of alternative groups that might carry them out. These alternative groups include:
(i) Scientific Committee sub-committees, workshops, special meetings;
(ii) research scientists attached to or formally employed by the Secretariat;
(iii) independent scientists or agencies working under contract or consultancy agreements;
(iv) scientists in national laboratories.

The Committee agreed that it was essential that work be commenced on the Comprehensive Assessment as soon as possible. Accordingly it recommended that the preparation of an inventory as outlined in step 1 of the Outline Work Plan for the Scientific Committee be carried out by the Secretariat as a matter of urgency, with a view to presenting a preliminary inventory (comprising general answers to A1(i) and (ii), and a listing of data held by the

IWC computing service) at the 1986 Annual Meeting of the Committee. If this can be done, then the Committee will be in a position at that meeting to review the inventory and start work on steps 2(i) and 3(i). The complete listing of data held by national governments and other research establishments is clearly a large task, particularly for long data series. In view of this, the Committee should undertake a final review of the requirements given in Annex E, at its next meeting, before asking the Secretary to send out formal requests for this information.

With these points in mind, the Committee developed the provisional timetable shown in Annex F

### 10.5 Costs

Carrying out the Comprehensive Assessment will make additional demands on the Secretariat and member governments. Some of these were discussed in section 10.4 above. It is too early to translate these demands into precise budgetary figures, especially for the later years of the assessment, but it is clear that provision will have to be made for some or all of the following:
(i) additional staff;
(ii) computing facilities;
(iii) workshops or other meetings;
(iv) contracts or other arrangements for special studies;
(v) new data collection;
(vi) national research activities.
(i) Implementing the Comprehensive Assessment will be a complex task, involving the interaction of at least three separate activities - study of management practices, the collection and compilation of data, and the methodology of assessment. The Committee's work might be most effectively facilitated by the provision of scientists whose time is dedicated totally to the undertaking of these tasks. It is envisioned that their work would be planned and supervised by the Committee. One possibility would be for the group to be attached to the Secretariat for the duration of the Comprehensive Assessment. Appropriate personnel might be found through normal hiring procedures or by secondment from interested member governments or other agencies. Typically costs for hiring (including salary, overheads and travel) such a scientist at international scales might range between $£ 30,000-40,000$ annually. Some members thought the Commission might consider seeking outside funding as an alternative to including this Item in its regular budget. Others did not believe this to be appropriate.
(ii) As noted in section 10.3 above, the Secretariat has made significant strides in entering relevant data into the computer but additional work remains to be done. Other computing tasks will no doubt emerge as the Committee proceeds with the Comprehensive Assessment. As in recent years the Secretariat will have to be augmented (e.g. by additional temporary staff) to undertake this task.
(iii) The Committee envisions the need to hold a series of workshops to examine and to resolve a number of methodological problems. For example, there is the need noted above to identify appropriate stock units and/or to examine the consequences to assessments and management advice of utilising erroneously defined stock boundaries. A dedicated staff group could assist the Committee's advance preparations for such workshops. Costs of such workshops would vary depending on location but might be kept to a minimum if held in Cambridge
where the database is located. Typical costs for such meetings might range up to $£ 5,000$ each (including provision for up to four invited participants).
(iv) As a result of workshops and other meetings undertaken as part of the Comprehensive Assessment it is likely that a number of follow-on studies will be defined. Most of these will likely be assigned to a dedicated staff group or else be taken on by national scientists. However, some studies will no doubt arise which require unique expertise or experience. For example, in recent years, Dr Butterworth has received contract support to prepare the IDCR minke whale cruise data for analysis by the Committee. Typical costs of such contracts range up to £10,000.
(v) Although the Comprehensive Assessment will analyse or re-analyse existing data, the Committee believes that it may be necessary to continue and/or initiate a number of data collecting activities. For example, the Committee may recommend continuation of the series of IDCR minke whale cruises, which has annual costs to the Commission typically ranging up to $£ 50,000$ although these are contingent on the cruise programme which costs national governments in excess of $£ 1,000,000$.
(vi) The Committee has assumed that all national research programmes will at least continue at their present levels, if indeed they are not enhanced. The continuing contribution of such efforts is viewed as a major component of the Comprehensive Assessment. All other needs identified above are viewed as enhancements which are necessary to accomplish the task in the time allotted.

The budget outline in Table 1 defines the scope of the needs for undertaking the Comprehensive Assessment. This indicates, for each of the Items, the estimated annual needs. It must be emphasised that at this stage the Committee is not recommending that provision be made for each of them. Rather, it wishes merely to indicate the overall scope of the task of the Comprehensive Assessment and the likely budgetary implications should any of the possibilities listed be required. A more detailed budget will be prepared after completion of step 1 of the Work Plan, and further amendments can be expected as the Comprehensive Assessment proceeds.

## 11. ADOPTION OF THE REPORT

Before closing the meeting, the Chairman, on behalf of the Committee, gave a vote of thanks to Charles Free, who has in his six years at the Secretariat established its Computing Facility. A considerable amount of essential data has now been encoded and verified and a most efficient systen developed for future encoding. The Committee has also made tremendous demands on the facility during its meetings and despite the lack of sleep (and often knowledge of the use, if any, the Committee would make of the results), Charles always managed to produce both results and a smile. The Committee wishes him well for the future. The Committee also thanked the Secretariat, and particularly Mrs V. Hunter who acted as meeting secretary, for the cheerful support given the meeting.

## 12. REFERENCES

Aguilar, A. and Nadal, J. 1984. Obtención de biopsias hipodérmicas en cetaceos. Inv. Pesq 48(1): 23-9.

Bushuev, S.G. and Ivashin, M.V. 1986. Variation of colouration of Antarctic minke whales. Rep. int. Whal. Commn 36: 193-200.
de la Mare, W.K. 1986. Simulation studies on management procedures. Rep. int. Whal. Commn 36: 429-50.
Fujino, K. 1960. Immunogenetic and marking approaches to identifying subpopulations of the North Pacific whales. Sci. Rep. Whales. Res. Inst., Tokyo 15: 85-142.
Ichihara, T. 1966. The pygmy blue whale, Balaenoptera musculus brevicauda, a new subspecies from the Antarctic. pp. 79-113. In: K.S. Norris (ed.) Whales, Dolphins and Porpoises. Univ. Calif. Press, Berkeley and Los Angeles. i-xv +789 pp.

Schnell, G.D., Douglas, M.E. and Hough, D.J. 1985. Further studies of morphological differentiation in spotted and spinner dolphins (Stenella attenuata and S. longirostris) from the eastern tropical Pacific. Southwest Fisheries Center Admin. Report LJ-85-04C. 30pp.
Wada, S. 1983. Genetic structure and taxanomic status of minke whales in the coastal waters of Japan. Rep. int. Whal. Commn 33: 361-3.
Wada, S. 1984. A note on the gene frequency differences between minke whales from Korean and Japanese coastal waters. Rep. int. Whal. Commn 34: 345-7.

## Annex A

List of Participants

| ANTIGUA \& BARBUDA | JAPAN | SEYCHELLES | IATTC |
| :---: | :---: | :---: | :---: |
| R. Payne | A. Mae | S.J. Holt | S.T. Buckland |
|  | S. Misaki |  |  |
|  | F. Nagasaki | SPAIN |  |
| AUSTRALIA | S. Ohsumi | L. Jover | IUCN |
| G.P. Kirkwood W.K. de la Mare | K. Shima | S. Lens | A. Rosenberg |
|  | S. Tanaka <br> K. Yamamura | H. Quiroga |  |
|  |  | USSR |  |
| DENMARK | MEXICO | M.V. Ivashin | INVITED PARTICIPANTS |
| F. Larsen | L.A. Fleischer | G. Kisseleva | P.B. Best |
|  |  | UK | J.G. Cooke |
|  | NETHERLANDS | P S Hammond | J. Gordon |
| FEDERAL REPUBLIC | K. Lankester | P.S. Hammond | J.A. Gulland |
| OF GERMANY | K. Lankester | J. Harwood A.R. Hiby | E.F. Harding |
| P. Deimer | NORWAY | J.W. Horwood |  |
|  | T. Oritsland | USA | IWC SECRETARIAT |
| ICELAND |  | R.L. Brownell | C. Allison |
| T. Gunnlaugsson | SAINT LUCIA | D.G. Chapman | C.A. Free |
| J. Sigurjonsson | F.J. Palacio | M.F. Tillman | R. Gambell |

## Annex B

## Agenda

1. Chairman's welcome and opening remarks
2. Appointment of rapporteurs
3. Adoption of Agenda
4. Arrangements for meeting
5. Review of available documents and reports
6. Definition of a comprehensive assessment
6.1 Eleınents of a comprehensive assessment
6.2 Relationship between management policies and advice required
6.3 Information required for a comprehensive assessment
7. Review of current methods for stock assessment and provision of advice on catch limits
7.1 Stock/management units
7.2 Available data
7.3 Assessment techniques
7.4 Additional reviews/studies of existing data or techniques required
8. Exploration of new management regimes
9. New information needed for a comprehensive assessment
9.1 New data
9.2 Alternative assessment techniques
10. Planning for the Comprehensive Assessment
10.1 Priority stocks
10.2 Priorities for reviews of existing information and collection of new information
10.3 Computing needs
10.4 Timetable
10.5 Costs
11. Adoption of report

## Annex C

List of Documents

## SC/A86/CA

1 CHAPMAN, D.G. Planning for the Comprehensive Assessment of Whale Stocks.
2 GULLAND, J.A. Notes on a Comprehensive Assessment.
3 HOLT, S.J. Planning a Comprehensive Assessment of Whale Stocks.
4 IKEDA, I. On Comprehensive Assessment - a Case Study of Antarctic Minke Whale.
5 Working Paper on the Comprehensive Assessment of Whale Stocks prepared by Japan (revised).

6 TANAKA, S. On a Practical Method for Stock Management.
7 LANKESTER, K. The History of MSY and its Consequences in Baleen Whale Management.
8 de la MARE, W.K. Elements of a Comprehensive Assessment of Whale Stocks - a Discussion Paper.
9 COOKE, J.G. The Assessment and Management of Whale Stocks.
10 BEST, P.B. The Comprehensive Assessment of Whale Stocks - Some Personal Opinions.

## Annex D

# Proposal for a Simulation Study on the Effects of Discrepancies between Stock Units and Management Units 

D.G. Chapman

It is agreed that many of the current whale management units do not contain a unique isolated genetic stock. It is possible that a single genetic stock may extend over two or more management units. On the other hand a single management unit may contain all or parts of several genetic stocks.

We also need to distinguish between breeding aggregations and genetic stocks. A genetic stock may be composed of several breeding aggregations with various degrees of mixing between them. A further complication is that many whale stocks have feeding areas which are distinct from the breeding areas and the relationship of different breeding aggregations and of the genetic stock they compose, to different feeding areas is often unclear Certainly it may be both complex and variable.

What is important from a management view is what is the possible impact of catches on this complex population structure? Can a management regime that appears to be rational and conservative lead to overexploitation or even
extinction of some of the genetic stocks within the management area? A full understanding of the situation probably requires knowledge of the interaction of the genetic stocks and also of the selection probabilities of the catching process with respect to the stocks.

In the absence of such detailed information, a simulation study may explore a range of possible outcomes of exploitation on such mixed genetic stocks within a management unit.

The aim of such a simulation study will be to calculate the probabilities of the exploitation reducing one of the genetic stocks to various levels (including zero) under a range of catches for the whole unit. Parameters that might be considered are the levels of catches (as a fraction of the total stock size), the period in years of the catches, the selection probabilities for the several stocks. Some restrictions in the choice of parameters will probably have to be made to reduce the number of calculations or simulations to a manageable level.

## PROPOSED DATA INVENTORY PER SPECIES/MANAGEMENT UNIT

1. Catch data
(a) Individual catch data ( $\min =$ species, length, sex, date, position if possible).
(b) Gross catch data (min $=$ number of whales, season/year).
(i) Which years/seasons available?
(ii) In what format?
(vii) Who holds data and restrictions on access?

## 2. Effort data

(a) Gross effort data (Net catcher day, Gross catcher day)
(b) Refined effort data (Corrected for time budget)
(c) Correction factors (Tonnage, horse power, ASDIC, weather, spotter aircraft, tow boats).
(i) Which years/seasons available?
(ii) In what format?
(iii) Who holds data and restrictions on access?

## 3. Age data

(a) Absolute age data (teeth/earplug laminae/tympanic bullae)
(b) Relative age data (ovarian corpora)
(i) Which years/seasons available?
(ii) What proportion of catch sampled each year/season?
(iii) What proportion of sample read each year/season?
(iv) Were transition phase data collected (each year/season)?
(v) Number of readers and readings and any indication of the reliability of readings
(vi) In what format?
(vii) Who holds data and restrictions on access?
4. Sighting data (surveys with associated sighting effort)
(i) Which years/seasons available?
(ii) Suitable for relative/absolute abundance indices?
(iii) In what format?
(iv) Who holds data and restrictions on access?

## 5. Marking data

(a) Natural marking data
(i) Which years/seasons available?
(ii) How many individuals identified?
(iii) In what format?
(iv) Who holds data and restrictions on access?
(b) Artificial marking data ('Discovery' tags)
(i) Which year/seasons avaiable?
(ii) How many recoveries made?
(iii) In what format?
(iv) Who holds data and restrictions on access?
6. Pregnancy rate data (biologists records only)
(i) Which years/seasons available?
(ii) Number of mature females examined each year/ season
7. Age at maturity data (from biological examination)
(i) Which years/seasons available?
(ii) Number of animals for which gonads examined each year/season (males and females separately)
8. Calving interval (from individually identified animals)
(i) Number of records available
(ii) Who holds data and restrictions on access?
9. Age at first calving (from individually identified animals)
(i) Number of records available
(ii) Who holds data and restrictions on access?

## Annex $F$

## Draft Timetable Workplan

(1) 1st round inventory (Secretariat/sub-Committee chairmen) - by 1986 meeting
(2) Examine methodological problems
(i) Identify priorities (Scientific Committee) - at 1986 meeting
(ii) Hold workshops, carry out simulations etc 1986/87
(3) Examine data quality
(i) Decide on actions (Scientific Committee) - at 1986 meeting
(ii) Collect new data - 1986 onwards
(iii) Make existing data more accessible (Secretariat) - 1986/1987
(4) 2nd round inventory: improved assessments
(i) Identify priority stocks for specific questions (Scientific Committee) - 1987 or 1988 meeting
(ii) Up-date/review assessment - 1987/1989
(5) Examine general features of whale dynamics
(i) Identify priority subjects, and methods of tackling them
(Scientific Committee) - 1988 or 1989 meeting
(ii) Hold workshops; carry out studies - 1989/1990
(6) Revise assessments

Complete 3rd round inventory - 1989/1990
(7) Present "comprehensive assessment" (Scientific Committee)- 1990 meeting.

# Comprehensive Assessment Workshop on Catch Per Unit Effort (CPUE) 

The Workshop was held at the Marine Research Institute, Reykjavik, from 16-20 March 1987. The director of the Institute, Jakob Jakobsson, welcomed participants to the meeting. A list of participants is given as Annex A.

## 1. CONVENOR'S OPENING REMARKS

Kirkwood described the background to the Workshop. Its terms of reference, determined at the 1986 Annual Scientific Committee Meeting, are given below.

Review the use of CPUE and other relative abundance indices for identifying and estimating trends in population abundance and productivity. This review should include but not necessarily be restricted to: detailed examination of case studies of CPUE or relative abundance series and associated operational data; simulation studies of refined or coarse CPUE series under various assumptions on behaviour of whales and whalers; modelling of CPUE series to examine the questions of censoring and outliers.

## 2. ELECTION OF CHAIRMAN, RAPPORTEURS

Kirkwood was elected Chairman. Donovan was appointed rapporteur with assistance from de la Mare and Butterworth.

## 3. ADOPTION OF AGENDA

The Workshop agreed that although the terms of reference given above referred to CPUE and other relative abundance indices, it would be appropriate to concentrate on CPUE. The adopted Agenda is given as Annex B.

## 4. ARRANGEMENTS FOR MEETING

The Workshop agreed to a work schedule suggested by the Chairman. He also stressed the need for participants to remember that, for many members, English was not their primary language. Stefánsson described the computing facilities available: an HP 9000/550 Unix computer, with several terminals, a laser printer and an 8 pen plotter; the computer was connected to the public data network and a telephone line so it was possible to send or receive mail and data from other computers.

## 5. REVIEW OF DOCUMENTS, DATA AVAILABILITY

The list of documents presented is given as Annex C. It was also agreed to request the rapporteur to draw up a preliminary listing of relevant literature to be discussed further at the Scientific Committee Meeting.

At this meeting the following data were available on the computing system:
(1) Icelandic catch and effort data, 1948-86;
(2) Japanese Antarctic minke whale catch and effort data 1971/72-82/83.

In addition, it was possible to access the International Whaling Commission (IWC) data base in Cambridge. The statistical packages GLIM and BMDP were also available.

## 6. TERMINOLOGY

The Workshop agreed it would be useful to develop a list and define terms traditionally used in the CPUE literature within the IWC. Donovan agreed to present such a list at the forthcoming meeting.

## 7. GENERAL CONSIDERATIONS OF CPUE AS AN INDEX OF ABUNDANCE

The fundamental question with respect to CPUE data is the extent to which it can be predictably related to stock abundance. The problem can be broken down into two questions.
(1) Is there a usable relationship between CPUE (however the effort is computed) and the abundance of animals in the area over which the whaling has occured?
(2) Is it possible to use this relationship to obtain an index of abundance for the total stock? This involves two stages: first, using the data for the area operated in each season to obtain an index for the total whaling ground (i.e. the area where whales might be expected to occur during the whaling season); second, to use this index to obtain an index of total stock abundance.

The IWC, in common with many fisheries bodies, has made extensive use of CPUE data. SC/M87/C1 and C3 reviewed the difficulties which the Scientific Committee has encountered over the years in the interpretation of CPUE data when used for stock assessment. These difficulties over the appropriateness of the use of such data have ranged from the belief that the series do not adequately take into account known operational or environmental factors, to problems arising out of inconsistencies between series thought to apply to the same stock, to the failure of series to provide a good fit to an estimated population trajectory. These problems, in addition to the questions of the relationship between CPUE and stock abundance, are addressed below.

### 7.1 CPUE as an index of local density

Several theoretical models have been developed for examining the relationship between CPUE and abundance, or for using CPUE data to estimate abundance
given certain major assumptions (in using CPUE data to assess the status of whale populations it has usually been assumed that searching is random with respect to the distribution of whales and that CPUE is proportional to density and in turn total exploitable population size).

SC/M87/C2 examined two aspects of modelling catch, effort and population size. First the author described the classic deterministic differential equation models of fishing. He then went on to develop and analyse some simple stochastic models which described the time trajectory for the catch frequency distribution. Finally he adapted the methods of renewal theory for a stochastic model of the catch process in which there is a handling time associated with each catch.

In discussions concerning the effects of non-random searching by catcher boats, the author emphasised that the methods of renewal theory could be used to model a catch process even when non-random searching occurred. The major prerequisite for development of such models is knowledge of the probability distribution functions for the time spent searching for a whale or pod of whales and for the time spent chasing and handling a whale.

The models in SC/M87/C2 assumed that the only form of variability in the data arise from sampling. However, analyses of whale CPUE series suggest that the variability in the data does not arise purely from sampling but also includes a large component arising from year to year variability in the catchability coefficient (de la Mare, 1984, Rep. int. Whal. Commn 34: 655-62; de la Mare, 1986, ibid 36: 419-23). The author stated that the results of SC/M87/C2 are preliminary and do not provide a satisfactory exploration of the data considered.

SC/M87/C4 presented some very simple and idealised models of non-random searching for whales which themselves have a non-random distribution. If the search process is structured, in the sense that there is a distinction between searching for patches of high density and searching within such patches, then the relationship between the encounter rate and the abundance of whales may not be linear. It will depend on the nature of the aggregations in the distribution of whales and how these change as overall abundance changes (i.e. to what extent are changes in overall abundance reflected in (1) frequency of patches, (2) area of patches, (3) density within patches, (4) range of the stock, (5) combinations of (1)-(4)), and on the methods of searching for such aggregations.

According to which model applies, the relationship between CPUE and abundance may be non-linear, or CPUE may be virtually independant of abundance over a wide range of abundance levels. These factors are additional to, and also interact with, those identified by Cooke (1985, Rep. int. Whal. Commn 35: 511-19).

The possibility was raised that, if a sufficiently small area was considered, the searching within that area could be assumed to be random (see Item 7.5). However, there may be circumstances where even this assumption does not hold.

The interpretation of CPUE requires a detailed understanding, both in qualitative and quantitative terms, of the methods used to search for and catch whales. In particular, the failure of the assumption that searching for whales is random with respect to the distribution of whales and that the whales themselves are independently distributed, has potentially major effects on the relationship between CPUE and abundance. The factors which may play a key role in determining the relationship
between CPUE and density can be classified in terms of the following: (1) deciding which area to search; (2) nature of searching within a locality; (3) relationship between numbers of whales seen and those actually caught; (4) factors relating to the distribution (patchiness) and behaviour of whales; and (5) exogenous factors (e.g. environmental conditions).

Some factors may affect the form of the relationship between catch, effort and abundance and hence need to be analysed even if they can be assumed to remain constant over time. For other factors that affect the ratio of CPUE to abundance, only changes need to be recorded.

Some aspects of the relationship arise from the consideration that searching within a locality may be non-random, i.e. whalers take advantage of the non-random distribution of whales to maximise searching efficiency. This can yield a non-linear relationship between the encounter rate per searching hour and the average number of whales per unit area (whatever size of area is specified), depending on the nature of the distribution of the whales and how this relates to average whale density. SC/M87/C3 gives some hypothetical examples of how such non-linearities can arise.

In order to relate catch and effort data to local whale abundance, at least a qualitative understanding of the method of searching is required, i.e.:
(1) how is search initiated;
(2) type of search (parallel, radial, sawtooth, etc);
(3) manner in which search pattern is modified following a sighting, catch, other cue, or receipt of information from another vessel;
(4) adjustment of search pattern according to operational factors e.g. remaining time available in day/trip, distance from factory ship/land station, instructions from factory ship/land station (e.g. whether processors busy);
(5) full data on the time and location of each catch, sighting, other cue and information received.

Information on other factors affecting searching efficiency is also required in the event that these factors are liable to change.

The relationship between the number of whales seen and those actually caught may not be one of simple proportionality (cf. SC/37/Mi11). Especially in areas of high density, a number of whales may come into view of which the number caught will be limited by the catching ability of the vessels, regardless of the number of whales seen; in this case some indices of catch per unit effort will be little related to abundance.

The following data may help to indicate the nature of the relationship: time budget data for each catcher, i.e. time each activity started and stopped (searching, confirming, chasing, handling); criteria for selection of whales to catch; group size; when to abandon chase; behavioural response of whale to whaling operation; all relevant technological and human factors, in so far as these are liable to change (a list of some of these variables is given in Appendix A of recent Schedules) ${ }^{1}$.

Where scouting boats are used in an operation, the success of catchers also depends on scouting information received. The following data may help determine the

[^2]contribution of scouting boats to the total searching effort: number, positions, time, species group size (and other characteristics, if applicable, of whales seen), time budget and other data (e.g. similar to IDCR records), cruise track design, information relayed to catchers and factory.

The nature of whale aggregations and the manner in which these change with overall abundance have implications for the relationship between CPUE and abundance. Some of the effect may arise directly as an effect on density, and some indirectly through the proportion of whales seen that are caught. For example, if the region occupied by a whale stock contracts to a favoured area, then the density within that area may tend to be independent of stock abundance. As another example, different effects arise if the density of schools remain constant but the number of whales within schools changes with changes in abundance (SC/M87/C4).

Some inference on these types of effects will require observations on the behaviour of whales, such as: school size; independent observation of distribution and abundance; tracking experiments (e.g. radio tagging); individual identification techniques.

### 7.2 Apportionment of effort in multi-species fisheries

It is not uncommon for whaling operations to take more than one species of whale in a season, e.g. Antarctic pelagic whaling for blue, fin and sei whales up to the 1960s, the Norwegian small-type operations for minke, bottlenose, killer and pilot whales. In such cases there is a problem in deciding on apportioning of effort. Attempts at overcoming this problem have been based on knowledge of the operations of the individual fisheries. For example in certain fisheries it may be possible to subdivide the data by time, by area, or by subsets of the fleets, in such a way that the appropriate effort can be assigned. Thus in the 1960s Antarctic operations, sei whales were caught early in the season and further to the north than fin whales (see Item 8). In situations where one species is the preferred species and the other effectively a by-catch, then it may be sufficient to subtract the handling time for the by-catch species from the total effort to deduce the effort for the preferred species. It should be remembered however that this may introduce some bias as the 'lost' time will be in an area of low density for the preferred species (hence the decision to take the secondary species).

In looking at series of CPUE data for multi-species fisheries, it is not sufficient only to determine the factors which make a gunner decide to chase a particular species but it is necessary also to examine whether the selectivity has varied over time (i.e. between seasons). This is similar to the situation in single species fisheries where selectivity for a certain segment of the population (e.g. larger animals) may vary with time or density.

### 7.3 Combination of different effort series

Several situations where one might wish to consider combining effort series were considered.

The first concerned the situation where more than one series exist which apparently relate to the same stock in the same area, but which show different trends (e.g. data for different pelagic fleets). Before considering this in more detail, the Workshop addressed the problem of how to compare two series. It was suggested that comparing the slopes of linear regressions against time was not appropriate since linearity may be in question and that it might be better to employ a non-parametric rank correlation test.

On some previous occasions where the Scientific Committee has been faced with two series showing different trends, it has selected one as the 'better' series, without adequately explaining the criteria used to make the decision. It was agreed that before discarding any series, efforts should be made to attempt to explain the differences. A close examination of operational factors might explain differences between series, for example length selectivity may vary so that the CPUE values relate to different segments of the population.

In previous assessments where series have been considered compatible, two approaches have been used to combine them: one has been to take a weighted average of the indices as a single series; the other has been to fit a population model simultaneously to both series allowing the fitting procedure to estimate the relative catchability coefficients and the population trend

De la Mare (1984, op. cit. and 1986, op. cit.) has shown that the between season variance of whale CPUE series, when expressed as a coefficient of variation, seems to be independent of sample averages (i.e. catch size). If this applies also to within-season variation, then simple averages rather than averages with weights which are functions of catch size would be appropriate.

In certain situations, a short recent series of detailed effort data are available in conjunction with a longer series of coarse effort data. Attempts have been made to compare these series in the overlapping time period to see if this can be used to obtain an 'improved' long time series. The Workshop agreed that all of these matters warranted further consideration.

### 7.4 Detection of outliers and effect of censoring

In past attempts to estimate temporal trends in CPUE series, the Scientific Committee has from time to time faced the problem that certain points are suggested to be outliers. This raises two questions: what test (or tests) should be used to determine whether a point is an outlier; and if an outlying point is so detected, should such point be omitted when the trend in the series is estimated?

The Workshop had insufficient time to consider this matter in detail. The following general principles were however suggested.
(a) When carrying out regressions to estimate temporal trends, it is advisable to test for 'influential' points (members of the data series to which the estimate of the trend is particularly sensitive). If there are such points, they should be more carefully scrutinised in respect of their reliability. A statistic that can be used to test for influencial points in multiple linear regression is Cook's distance (R.D. Cook, 1977. Detection of influential observations in linear regression. Technometrics 19: 15-18).
(b) The presence of outliers may lead to erroneous inferences regarding trends. Various statistics which warn of the presence of outliers have been proposed in the statistics literature. When the presence of outliers is suspected, the data should be re-examined in respect of the possibility that further explanatory variables are required for an adequate description.
(c) Standard methods exist for the detection of a single outlier in a series, but these are based on the assumption that residuals are normally distributed. More recent developments, based on the bootstrap method (e.g. R.S. Sparks, 1988 (In Press). A distribution-free method of detecting outliers in regression. Commun. Statist. A17(3).), do not require the assumption of normality, and
have the ability to detect multiple outliers. For CPUE data, the assumption of normality of residuals would seem unlikely to be appropriate. Further the distribution of the CPUE statistic will not necessarily be the same from year to year; for example plausible mechanisms can be suggested that could lead either to an increase or a decrease in the variance of CPUE as a function of density. Such circumstances preclude the use even of the distribution-free procedures referred to above, as a rigorous statistical basis for outlier detection.
The Workshop did not discuss the question of censoring.

### 7.5 Relationship between density on whaling grounds and population abundance

Most of the discussion under this heading was related to the case of the Southern Hemisphere minke whale. Many of the points raised have relevance to other operations, but equally there are important additional problems with some other fisheries, such as those operating from land stations. There was not sufficient time to consider the specific problems of abundance estimation for CPUE for other operations in depth.

Fig. 1 (from SC/37/Mi11 revised) shows that the Southern Hemisphere minke whaling operations from 1971/72 to 1982/83 were restricted each year to a narrow band (possibly with some gaps) just to the north of the ice-edge. It was suggested that density indices be obtained from CPUE data for suitably small longitudinal strata within these bands and that such stratification may largely remove any problems of CPUE interpretation associated with the non-random-search nature of the catching operation. Local whale movement was also cited as a reason that such non-randomness was likely to be of little consequence. An examination of the distribution of search intervals may provide some insight into this matter. However, it was also suggested that the assumption that the searching operation tended towards randomness in the limit of stratification into small areas, may not be justified.

Given circumstances in which the assumption of random searching can be justified, density indices for longitudinal strata within a band could be calculated between the longitudinal limits for the management area in question, to provide an index of the abundance of whales on the whaling grounds for the area and season concerned. A problem that arises however, is that in some seasons the band in which whaling took place does not extend to the longitudinal limits of the management area. A solution suggested was to estimate the average relative density values for longitudinal strata within the band using regression methods. This exercise would provide comparable annual abundance estimates for the whaling grounds within a management area; however, it relies on the untested asssumption that the pattern of density variation with longitude in an area remains fixed as total abundance changes. Suggestions were made that attempts be instituted to link such possible density variation patterns to environmental factors.

An important problem raised was the fact that the whales within a management area are not fixed in position throughout the season, but follow migration patterns which may vary from year to year. It was questioned whether it was valid to connect CPUE density indices to abundance values without a more complete understanding of this process. Partial account could be taken of this aspect by incorporating parameters allowing for monthly
variation of density in the whaling grounds (a consequence of migration patterns) in the regression procedure described in the preceding paragraph. It was further suggested that the magnitude of biases that might arise in estimates of abundance due to incomplete account being taken of migration effects, be investigated using simulation models.

CPUE data from whaling operations in a narrow band in the vicinity of the Antarctic ice edge provide no information on the abundance of whales further north in the management area. The Workshop suggested a stratification approach be investigated, with the relationship between abundance of whales in the whole area, and that on the whaling grounds, being inferred from other information including the IDCR sighting survey programme and Japanese scouting vessel data. Such an approach assumes that the latitudinal density pattern does not change with total abundance. If CPUE data are considered for use in the future, this assumption should be checked at regular intervals by independent survey, as typical commercial operations would not provide the required data. This is important because a drop in abundance could be reflected by a decrease in the geographic extent of the whale distribution without necessarily any change in the density on the whaling grounds (see Winters and Wheeler, 1985, Interaction between stock area, stock abundance and catchability coefficient. Can. J. Fish. Aquat. Sci. 42: 989-98).

## 8. CASE STUDIES

While it was not possible during the Workshop to discuss any particular CPUE series in detail, it is apparent from the general discussion that knowledge of operational factors relevant to each series is of vital importance to their interpretation. This was brought out by the insights provided to discussions by the description of Japanese pelagic minke whaling operations given by Yamamura and summarised below.

Overall fleet strategy is determined in Japan before the fleet moves south and is based on two factors: the quotas allocated and the optimum duration of the season. The latter is decided in the light of legal considerations, processing capacity, behaviour of whales, expected weather/sea conditions and additional duties to whaling. As the employees are paid throughout the year, there is no economic pressure in this regard to make the season as short as possible. Based on the quota and expected length of the season, in 1986/7, the required average catch was about 20 animals per day. Allowing for variations in density of whales, their distribution and the weather, a maximum processing capacity of 26 whales per day was set (requiring a non-operational crew of 160 on the factory ship). In earlier years with a larger quota and/or two other species under exploitation, the processing capacity was up to 100 whales.

The general cruise track is designed to waste as little time as possible in transit. Operations begin in Area IV (the catchers arrive and search in the Area two days before the factory ship) and the fleet moves first to the west, and then to the east, so as to arrive as close to the boundary when completing the Areal quota to reduce transit time. A similar consideration applies in Area V and then Area VI, with the aim being to continue moving eastwards through
these Areas, so that when the total quota is reached the vessels have the shortest possible transit time back to Japan.

On a smaller scale the predominant factor determining strategy is the expectation of catching 26 whales per day. Thus the aim is not necessarily to go to areas with the highest density, but just areas where it is likely this operational target can be met. Early in the morning ( 6 am ), all catcher boats begin co-operative searching, moving east or west along the ice edge, about 3 miles apart. If one catcher finds a group of whales, a decision is made based on the school size and occurence of other vessels' sightings as to whether the other vessels should join it. In very dense areas, the catchers may select for larger whales provided the operational target of 26 whales per day can be met (in any event whales below 26 ft are not taken). In addition, there is a desire to spread out catch times such that the whales are processed on the factory ship in as fresh a state as possible. Thus, if several whales are caught early in the morning, catching may be halted even if whales are in the vicinity, although searching may continue.

There was a brief discussion as to how procedures had changed over time. In terms of minke whaling perhaps the largest change was in the processing capacity mentioned earlier. In terms of all Antarctic pelagic whaling the most important changes affecting operational strategy were the changes from blue whale units and 'Olympic' competitive whaling to species and Area catch limits. When fin and sei whales could be caught under the blue whale unit system, selection for species was based largely on a time/region basis. At the beginning of the season, moving south, the whalers would encounter sei whales only, later they would move further south to regions in which fin whales were predominant although sei whales were occasionally found. In this regard it is interesting to note that under the BWU system based on oil yields, six sei whales were considered equal to one blue whale, whereas to the Japanese industry, which was largely directed towards meat, sei whales were relatively more valuable than blue whales in terms of product. This became a factor in determining preferred species after the fin whale quota was reduced.

The above description confirms the view that detailed descriptions of operational factors are essential to interpreting CPUE data, although it was noted that it is easier to obtain such information in cases such as the above, where a corporate strategy is involved, than in for example, small-type whaling operations, where each skipper develops his own strategy.

Yamamura, Sigurjónsson* and Øien agreed to attempt to provide papers detailing the operational factors involved in Japanese pelagic minke whaling, Icelandic coastal catching of large whales and Norwegian small-type whaling, for the next Scientific Committee Meeting. It was recognised, however, that even with this information,

[^3]considerable problems may remain in determining what relationship exists between CPUE and abundance (see Item 7.1).

## 9. RECOMMENDATIONS AND GUIDELINES FOR FUTURE USE OF CPUE DATA

This Workshop was called because CPUE data may contain important information but difficulties have been found in relating changes in CPUE to whale abundance. It has been possible to clarify a number of factors which are important in this process, but the precise relationships remain uncertain.

Despite these problems, because CPUE series represent the only data available for some stocks, attempts should be made to extract the maximum possible information from the data. The Workshop was unable to state what level of information will be obtainable in any particular case.

While, given the problems identified during the Workshop, some members believed that CPUE is not suitable as a primary means of assessing stocks and their trends except where that suitability can be demonstrated on a case by case basis, it was agreed, now that the problems have been identified, they should be examined, and efforts made to obtain other information necessary to compile valid indices of population abundance from catch and effort data.

The Workshop recommended that if the Scientific Committee decides to try to use CPUE in the Comprehensive Assessment for any stocks, then it should ask the Commission to request member nations who exploited those stocks to supply a detailed description of the methods and strategy of the operation (particularly with respect to any changes with time) and the other information suggested in Item 7.1. When this information is obtained, the Scientific Committee should develop models of the relationship between CPUE and abundance and carry out appropriate tests of the validity of the models through simulation or in other ways.

The possible implications of whale movement for the use of CPUE data to estimate trends in abundance was discussed under Item 7.5. The Workshop recommended that the Scientific Committee should consider attempting to develop models which could help in determining the effects of this on catching and CPUE and in identifying gaps in our knowledge of significance to the use of CPUE. Such a study should consider both large (migrational) and small scale movements.

In addition, it is recommended that the Scientific Committee consider the need for surveys and/or other field studies as a means to examine specific problems identified by this Workshop.

## 10. ADOPTION OF REPORT

The Workshop wished to thank the Marine Research Institute for the efficient way in which it had organised the meeting.

## Annex A <br> List of Participants

| DENMARK | T. Miyashita | SEYCHELLES | IWC |
| :--- | :--- | :--- | :--- |
| F. Larsen | T. Nakamura | S. Holt | G. Donovan |
|  | S. Ohsumi |  | R. Gambell |
| ICELAND | K. Shima |  |  |
| T. Gunnlaugsson | K. Sakuramoto | SPAIN | INVITED PARTICIPANTS |
| K. Magnússon | K. Yamamura | H. Quiroga | D.S. Butterworth |
| J. Sigurjónsson |  | S. Lens | J.G. Cooke |
| G. Stefánsson <br> G. Vikingsson | NETHERLANDS |  | W.K. de la Mare |
| JAPAN | K. Lankester |  | G. Kirkwood |
| H. Kishino | NORWAY | USA | D.B. Sampson |
|  | N. Øien | D.G. Chapman | S. Zahl |

## Annex B <br> Agenda

1. Convenor's opening remarks
2. Election of Chairman, Rapporteurs
3. Adoption of Agenda
4. Arrangements for meeting
5. Review of documents, data availability
6. Terminology
7. General considerations of CPUE as an index of abundance
7.1 CPUE as an index of local density
7.2 Apportionment of effort in multi-species fisheries
7.3 Combination of different effort series
7.4 Detection of outliers and effect of censoring
7.5 Relationship between density on whaling grounds and population abundance
8. Case studies
9. Recommendations and guidelines for future use of CPUE data
10. Adoption of report

## Annex C List of Documents

SC/M87/C

1. LANKESTER, K. How suitable are whale CPUE data for population assessments?
2. SAMPSON, D.B. Some models of catch and effort.
3. COOKE, J.G. Impediments to the use of CPUE data for whale stock assessments.
4. COOKE, J.G. Further notes on the relationship between catch per unit effort and whale abundance.

# Comprehensive Assessment Workshop on Management 

The Workshop was held at the Marine Research Institute, Reykjavík from 23-25 March 1987. A list of participants is given in Annex A.

## 1. CONVENOR'S OPENING REMARKS

Kirkwood outlined the background to the Workshop. The terms of reference were to:
(1) examine the scientific aspects of recent simulation and theoretical studies of alternative feedback management strategies for whale stocks;
(2) develop recommendations for the nature and directions of further studies on this topic.

## 2. ELECTION OF CHAIRMAN AND RAPPORTEURS

Kirkwood was elected Chairman. Donovan was appointed rapporteur with assistance from various members of the Workshop (particularly Butterworth and de la Mare) as appropriate.

## 3. ADOPTION OF AGENDA

The Agenda adopted is given as Annex B.

## 4. ARRANGEMENTS FOR MEETING

The Workshop agreed to a work schedule suggested by the Chairman. He also stressed the need for participants to remember that for many members English was not their primary language.

## 5. REVIEW OF DOCUMENTS

Annex C lists the relevant papers presented to previous meetings and papers prepared especially for the Workshop.

## 6. CONSIDERATION OF SCIENTIFIC ASPECTS OF SIMULATION AND THEORETICAL STUDIES OF ALTERNATIVE MANAGEMENT STRATEGIES

The Workshop considered technical details of the papers presented at this and previous meetings. It agreed that it was appropriate for the authors to incorporate comments and clarify points in the revised versions of the papers submitted for publication. The report below (Items 6.1 and 6.2) therefore contains only brief summaries by each author of their papers, concentrating on assumptions, decision rules and data requirements, rather than detailed technical information.

### 6.1 Review of the new management procedure (NMP)

SC/M87/M2 gave an overview of recent studies by de la Mare pertinent to whale management, contained mostly in papers SC/37/O 14 and SC/38/O 3. Recent analyses of the properties of estimating the yield of a whale stock lead to the conclusion that such estimates are only sufficiently precise to indicate the general range in which the yield might be. Under this circumstance, and particularly in the light of uncertain estimates of stock size and trend, the properties of any management procedure need to be examined in terms of their ability to detect and correct errors made in setting catch limits.

Some properties of the NMP were examined by means of simulation trials in SC/37/O 14. In trials using CPUE data which were proportional to abundance, but with random error dependent on the catch size, it was found that there was a tendency for the NMP to lead to under-exploitation of the stocks. This problem was shown to be due to lack of independence between information (CPUE data) and control (catch limits). The net result of the trials was to show that the probability of reducing a stock to below the protection level was low, but at the cost of failing to attain productive use of the stock. Applying the NMP with unbiased independent estimates of abundance (e.g. from sightings surveys) solved the problem of drifting to under-exploitation, but also led to increased probabilities of reducing stocks to below the protection level. With either form of data, the NMP does display some adaptive behaviour, but the rate of adaptation is slow. Nonetheless, the NMP can be modestly successful, particularly if the degree of variability in the data is low and there is a degree of separation between information and control.

### 6.2 Alternative management strategies

### 6.2.1 De la Mare's studies

SC/37/O 14 and SC/38/O 3 included some tests on a revised management procedure which replaced explicit reference to MSY and MSYL with an arbitrary target level and protection level. The revised procedure incorporates explicit feedback by altering catch limits in direct proportion to the difference between the current population level and a target level. A fitted population model was used to estimate the degree of depletion. In most cases, catch limits were based on estimated replacement yields from the fitted models, modified according to the feedback rule.

In the revised procedure, the use of CPUE data was still found to have the problem of drifting to under-exploitation. With unbiased absolute abundance data, the overall performance of the revised procedure is improved in comparison with the NMP; the time span of adaptation is shorter and the probability of a stock being depleted to below the protection level is reduced.

However, if the absolute abundance estimates are biassed, the procedure will fail to stabilise the true population at the target level. This problem can be overcome by treating the data as a relative abundance index, but with an increase in the frequency with which stocks are estimated to require protection.

With the revised management procedure the use of certainty equivalence (acting on the estimates of decision parameters as if they were without error) was moderately successful in dealing with the poor precision of assessments from short data series, but led to very variable catch limits over the first 50 to 100 years of exploitation. Modifications to the decision rules are required if stable catch limits are an important objective. Using the average catch as input to the decision rule seems a promising line of enquiry. A further modification which took into account uncertainty in the estimate of depletion was also found to have some promise as a means of controlling the probability of depleting a stock to below the protection level.

The modifications offered to the NMP were not intended as concrete proposals for a revised management procedure at this stage, but were intended to illustrate that improvements could be made. Some principles for the design of improved management procedure were identified as: (1) incorporating explicit feedback in management decision rules; (2) framing these rules in terms of parameters which can be robustly estimated; and (3) achieving substantial independence between information and control. It was suggested that simulation procedures have an essential role in the further design of management procedures.

### 6.2.2 Cooke's approach

SC/A86/CA9 outlined a procedure based on using a provisional value for the MSY exploitation rate (see definition of MSY\% given under 6.4.1 (b)) along with a decision rule specifying what proportion of this rate can be taken, according to the estimated degree of depletion of the stock. This exploitation rate is converted to a catch limit using an estimate of current population size. The estimate of current population size is obtained by averaging survey estimates of the population obtained over the most recent 10 years, after adjusting each annual estimate by the amount of subsequent catches. This filtered estimate is then adjusted for uncertainty by taking its lower $95 \%$ confidence limit. The depletion of the stock is estimated by fitting a population model through the estimate of current population size using the known catch history and the provisional value for the MSY exploitation rate.

The provisional value is updated only when the data indicate a value significantly different from the provisional value at the $95 \%$ level. The new value adopted is that value closest to the provisional value that does not differ significantly from the estimated value at the $95 \%$ level of significance. The new value is used both for calculating the population trajectory (to estimate stock depletion) and for specifying the MSY exploitation rate. The decision rule then determines what fraction of this rate can be taken at this level of stock depletion. The estimate of current population size converts this exploitation rate into a catch limit.

The aim of biassing the estimate of MSY exploitation rate towards the provisional value is to reduce fluctuations in catch limits and to minimise risks of excessive stock depletion.

Some single realisations of this procedure, given annual estimates of abundance, showed that it was able to adapt the rate of exploitation so that the population stabilised near the MSY level and that, in doing so, catch limits did not vary greatly from year to year. However, performance in this latter regard was less satisfactory when the interval between abundance estimates was increased to five years.

### 6.2.3 Sakuramoto and Tanaka's approach

Sakuramoto and Tanaka $(1986,1987)$ conducted a simulation study for the case where Tanaka's method (1980) was adapted to apply to the Southern Hemisphere minke whale populations. The noteworthy feature is that assumptions about the type of population model (such as use of the Pella-Tomlinson model), MSY and MSYL are not needed for this management procedure. The only information needed to manage the population is a relative value (or 'index') of population abundance. (For the paragraphs immediately following, this index is taken to be linearly proportional to abundance; investigations where this is not so have also been carried out, and are discussed later in this section.)

The basis of this procedure is to set some value of this index as the target level (TL). The catch limit is then increased or decreased depending upon the present level of the index relative to the target level and the rate of change of the index. If the index of population size is $x \%$ larger than the TL, then the present catch limit is increased by $h x \%$. Further if the index is decreasing at a rate of $y \%$ of its current level per year, then the present catch limit is decreased by gy\%. Note therefore that this procedure sets the catch limit for the next year as a proportion of the value of the previous year's catch limit where this proportion depends on the two factors indicated.

Initially, Monte Carlo simulation was conducted for the case where TL, $g$ and $h$ are fixed, and the stability of the stock-harvesting system was investigated. Thus the minimum population level reached (which is related to the possibility of extinction of the population), mean catch limit, and similar statistics were examined. Further, an extension to the feedback procedure which adjusts TL (whose value was kept fixed in the initial simulations) towards the MSYL was discussed. Forecasts were then made for future values of the relative population abundance index; using these, a further modification of the procedure was introduced which allows the control variables $g$ and $h$ to be adjusted to make the system more stable, and the results of this were discussed for various cases.

Simulations were also conducted for the case where the stock boundary does not coincide with the management area. Finally, a simulation was conducted for the case where an abrupt change in population level occurred for some reason.

The conclusion from the above investigations is that when an index of relative abundance of the population is available, it is very possible that robust management can be achieved even if the index is subject to some measurement error. (The case of a uniform distribution of such measurement error over the range $-35 \%$ to $+35 \%$ was considered.) For the case where CPUE is used as the index of population abundance, but CPUE is proportional to the square root of population size (rather than linearly proportional as previously assumed) it is still possible to maintain robust management by increasing the control variable g , and setting TL higher. A highly flexible
management procedure can be obtained by incorporating the procedure for adjusting TL and $g$ into the control system. This management procedure can accommodate a possible change in an ecosystem's carrying capacity which might be caused by alteration in environmental factors, whether this changes the population's carrying capacity value in proportion, or alternatively leads to an adjustment in the relative carrying capacity values for different species inhabiting the ecosystem. The authors consider that this procedure to manage the stock shows considerable promise even when stock boundaries and management areas do not coincide (i.e. overlap each other), though only limited investigation of models that mimic such effects has been possible thus far. Even when an abrupt change of population size occurs, such as a sudden decline of $30 \%$, its impact on the performance of the procedure is virtually negligible. Further simulation and considerations more representative of the actual situation would be useful.

### 6.2.4 Magnusson's approach

In SC/M87/M4 a simple feedback system to regulate catches from a whale stock was considered. Catches were modified on the basis only of the slope of log CPUE regression against time. It was assumed that no information was available regarding a desirable target level. Two control parameters were available: the feedback gain, and the number of past years used in the regression. Preliminary deterministic simulations indicate that the stock will stabilise, provided that the feedback gain is greater than unity. The stock will stabilise at a higher level when either the feedback gain, or the number of past data points used in a regression, is increased.

### 6.3 Management procedures and stochastic control theory

 Following the presentation of the papers described in Items 6.1-6.2, Dr Jacobs of the Engineering Science Department, Oxford University, was asked for comments in relation to his expertise in control engineering. His detailed comments are given in Annex D but three main points were made.He noted that the general character of the previously described management systems was of an uncertain dynamic system under conventional feedback control together with a jacket monitoring the basic performance. He pointed out that:
(1) The dynamic system is non-linear, whereas the control strategies presented to the meeting were matched to linear systems. It is quite unlikely that a strategy suitable for a linear system would give the best regulation here.
(2) The feedback control should be explicitly recognised as being comprised of a control unit and an estimator: these have different roles and can be independently designed.
(3) The nature of the jacket (e.g. an imposed mechanism to prevent over-reaction to a controller) would be dependent on the particular problem. General control theory may not help in its design.

### 6.4 Evaluation of proposed management procedures

### 6.4.1 Initial screening procedure

The Workshop agreed that in order to evaluate, and in particular to compare, different proposed management procedures, a standard underlying population model should be used in simulation trials. Certain parameters of this model would be fixed. The trials would explore performance for various sets of values for the other
parameters, and the initial conditions. The following sections detail the agreements in these respects.
The Workshop recognised that the following trials constituted only an initial screening procedure for a proposed management procedure; further more stringent tests would subsequently need to be applied to procedures that showed promise following these initial trials (see Item 6.4.2).

## (A) Model and fixed parameters

The generalised production model (as specified by De la Mare in SC/37/O 10 and O 11) would be used to describe the underlying (or 'true') population dynamics of the exploitable population level (P) :

$$
\begin{align*}
P_{t+1}=\left(P_{t}-C_{t}\right) e^{-M}+ & \left(1-e^{-M}\right) P_{t}-t_{m} \\
& {\left[1+A\left\{1-\left(\frac{P_{t-t_{m}}}{K}\right)^{z}\right\}\right] } \tag{1}
\end{align*}
$$

The following parameters would be given fixed values for all trials:
Natural mortality rate $(\mathrm{M})=0.05 \mathrm{yrs}^{-1}$
$\begin{array}{ll}\text { Age at maturity }\left(\mathrm{t}_{\mathrm{m}}\right) & =7 \mathrm{yrs} \\ \text { Carrying capacity }(\mathrm{K}) & =10,000\end{array}$
The model assumes a gestation period of 1 year, so that the age at first parturition is 8 years, and further assumes that the age at recruitment is equal to the age at first parturition.

## (B) Simulation trials

This section considers indices of abundance obtained both from the catching operations (e.g. CPUE), and indices independent of the catching operations (e.g. independent sightings survey estimates of abundance). For convenience of expression, 'sightings abundance estimates' is used below to refer to the latter category.

In the list of trials that follow, the headings and symbols used have the following meanings.

MSYL. The ratio $\mathrm{P} / \mathrm{K}$ at which the model of equation (1) yields maximum sustainable yield. A desired ratio is achieved by adjusting the $z$ (density dependent exponent) parameter (e.g. for MSYL $=60 \%, z=2.39$ ).

MSY\%. The ratio of MSY to MSYL expressed as a percentage. A desired ratio is achieved by adjusting the A (resilience) parameter.
$\mathbf{P}_{\mathbf{o}}$. The true exploitable population level at the start of the year ( $t=0$ ) in which the (proposed) management procedure is first implemented.
J. A measured index of the exploitable population level $P$ (e.g. sightings survey abundance estimate or CPUE measure).
Error distribution. The distribution function for the random noise associated with a measurement of $J$. Two forms are used. The log-normal can apply either to a sightings survey abundance estimate (used as a relative index, or an unbiased absolute abundance estimate should the procedure require this) or to CPUE data, and has a constant coefficient of variation (CV). The gamma distribution is used as an alternative in respect of CPUE data only; the over-dispersion coefficient (u) of the distribution is chosen so that if a catch equal to the maximum sustainable yield were taken, the coefficient of variation would have the specified CV value.
CV. The coefficient of variation specified as above for the various error distributions.

The trials listed in Table 1 are all to be carried out for the case that the index J (which is measured every year) is linearly proportional to P : e.g. for the log-normal error distribution:

$$
\mathrm{J}=\mathrm{qPe} \mathrm{q}^{\varepsilon} ; \varepsilon \text { is from } \mathrm{N}\left[0, \sigma^{2}\right] \text { where }(\mathrm{CV})^{2}=\mathrm{e}^{\sigma^{2}}-1
$$

$\mathrm{q}=\mathrm{a}$ constant (the catchability coefficient)
For the gamma distribution for CPUE data, the probability density function for effort, E, for a given catch C and true exploitable population P is given by

$$
\begin{aligned}
& \mathrm{f}(\mathrm{E} \mid \mathrm{C}, \mathrm{P}, \mathrm{u})=\left[\mathrm{v}\left(\mathrm{vE}^{\mathrm{r}-1} \mathrm{e}^{-\mathrm{vE}}\right] / \Gamma(\mathrm{r}),\right. \\
& \text { where } \mathrm{v}=\frac{\mathrm{qP}}{\mathrm{u}}, \mathrm{r}=\frac{\mathrm{C}}{\mathrm{u}},(\mathrm{CV})^{2}=\frac{\mathrm{u}}{\mathrm{C}}
\end{aligned}
$$

The constant q applies to the (deterministic) relation CPUE $=q$ P.

Table 1
Evaluation of proposed management procedures: simulation trials

| Trial no. | P | MSY (\%) | MSYL (\%) | Error <br> distribution | CV |
| :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | K | 1 | 60 | log-normal | 0.4 |
| 2 | 0.3 K | 1 | 60 | log-normal | 0.4 |
| 3 | K | 4 | 60 | log-normal | 0.4 |
| 4 | 0.3 K | 4 | 60 | log-normal | 0.4 |
| 5 | K | 1 | 80 | lognormal | 0.4 |
| 6 | K | 4 | 80 | log-normal | 0.4 |
| 7 | K | 1 | 60 | ganma | 0.4 |
| 8 | 0.3 K | 1 | 60 | ganma | 0.4 |
| 9 | K | 1 | 60 | log-normal | 0.2 |
| 10 | 0.3 K | 1 | 60 | log-normal | 0.2 |

With respect to the trials, the following points should be noted.
(a) If the management procedure uses sightings survey data only, trials 7 and 8 may be omitted.
(b) If the management procedure uses CPUE data only, trials must also be carried out for the case that J is proportional to the square root of P : e.g. for the log-normal error distribution:

$$
\mathbf{J}=\tilde{q} \sqrt{\mathbf{P}}^{\varepsilon}
$$

For the gamma distribution in this case, $v=(\tilde{q} \sqrt{\mathrm{P}}) / \mathbf{u}$.
(c) If the management procedure uses sightings abundance estimates only, trials must also be carried out for the case that such estimates become available only at the start of the year $t=0$ and every fifth year thereafter, rather than every year (in either event, no sightings abundance estimate is available prior to $t=0$ ).
(d) For cases where $P_{o}=0.3 \mathrm{~K}$, there has been a constant catch over the last 30 years, commencing with $\mathrm{P}_{-30}=\mathrm{K}$. The magnitude of this constant catch is to be calculated deterministically using equation (1).
(e) For cases where $P_{o}=0.3 \mathrm{~K}$, and a management procedure that uses CPUE data, trials are to be carried out both assuming CPUE data is, and is not, available for the preceeding 30 years (i.e. all years $t=-30$ to $t=-1$ ).
(f) Insofar as a proposed management procedure may usefully assume certain information (e.g. if a population model is to be fitted to measured indices), the following may be assumed known by the managers:
(i) the generalised production model (as in equation (1)) for the population dynamics;
(ii) $\mathrm{M}=0.05$;
(iii) $\mathrm{t}_{\mathrm{m}}=7$;
(iv) for cases where $P_{o}=K$, it is known that the population was at carrying capacity at time $t=0$, but the true numerical value of $K$ is not known (e.g. for population model parameter estimation purposes);
(v) for cases where $P_{o}=K$, the constant catch level over the years $t=-30$ to $t=-1$ is assumed known, and it is also known that $\mathrm{P}_{-30}=\mathrm{K}$ in the same sense as in (iv) above.
(g) All trials are to assume that exactly the catch limit specified by a management procedure will be taken.
(h) Calculations are to be carried out using floating point arithmetic.

## (C) Specification of proposed management procedures

The nature of the procedure is entirely in the hands of its proposer. Its specification must take the form of a set of detailed and complete decision rules, information requirements, and, if appropriate, methods of parameter estimation. It is recognised that:
(i) some procedures might require both initial and long-term components, if the long-term component can only be implemented after indices became available from the initial years;
(ii) procedures utilising both CPUE and sightings abundance estimates are also possible, but under such circumstances appropriate adjustments need to be made to notes (b) and (c) of the previous section.
(D) Number of simulations and presentation of results*

In each case, the management procedure should first be simulated deterministically (zero noise in the abundance index J), and the population and catch trajectories for the first 200 years plotted. If the behaviour seems reasonable, 100 trials (with measurement noise), each of at least 100 years, should be simulated.

Note that the full set of statistics specified below should only be collected for the single set of control law parameters for which the proponents consider their procedure to provide the 'best' overall performance. Some indications of sensitivity to variations in the control law parameters should be reported; however, further histograms are not necessary for this purpose, and comparison of changes in median values as well as $90 \%$ ranges (i.e. centralised range including $90 \%$ of simulation run results) for the statistics of interest should be adequate.

Bearing the above in mind, the following statistics from each trial should be collected, so that the associated histograms can be plotted:
(i) lowest true population level;
(ii) true population after 10,30 , and 100 years;
(iii) cumulative catch after 10,30 and 100 years;
(iv) root-mean-square change in catch limit between adjacent years (a measure of catch stability) after 10, 30 and 100 years;
(v) time the stock is first protected (if appropriate to the procedure proposed);

[^4](vi) time catches from the stock first recommence (if initially protected in terms of the procedure proposed).
Further, time series of the true population and annual catch should be plotted for:
(i) Cases 1, 2, 3 and 4 of Table 1 with $\mathrm{CV}=0$ (i.e. no measurement noise) for a 200 year period.
(ii) One realisation of case 1 and case 9 of Table 1 with noise (i.e. as specified for those cases) for a period of 200 years.
The above represents the minimum set of presentations needed to be available for inspection if required; not all of these results need be included in a description of the results for general circulation. Authors may add presentations if they desire.

From these outputs the performance of various proposed management procedures can be compared in the context of the following three general aims, and the trade-offs between these.
(a) Stability of catch limits, which would be desirable for the orderly development of the whaling industry, and would be indexed by the root-mean-square inter-annual change in catch limits.
(b) Acceptable risk level that a stock not be depleted (at a certain level of probability) below some chosen level (e.g. some fraction of its carrying capacity), so that the risk of extinction of the stock is not seriously increased by exploitation. This would be indicated by the distributions of the true population level and the distribution of lowest points in the stock trajectories.
(c) Making possible the highest continuing yield from the stock, which would be shown by the distribution of cumulative catches.

In respect of point (a), it was also suggested that comparisons could usefully be made to the optima achievable in the circumstances of exact and complete information (e.g. for the case $\mathrm{P}_{\mathrm{o}}=\mathrm{K}$, this would correspond to an immediate catch to reduce the population to MSYL, followed by continuing catches of the maximum sustainable yield).

### 6.4.2 Second stage testing of selected procedures

Procedures which yield promising results in the above initial process will then require more stringent testing of their performance in the presence of adverse factors. These factors, which may vary on a case-by-case basis may include:
(1) erroneous stock identification;
(2) bias in population estimates;
(3) a wider range of underlying relationships between CPUE and abundance (if applicable);
(4) changes in age or length at recruitment (selectivity);
(5) long term changes in the yield curve (either as a result of a changing carrying capacity due to environmental changes, or changes in population parameters usually considered constant);
(6) erroneous assumptions regarding the error distribution of the abundance index;
(7) erroneous assumptions regarding the population model
(e.g. shape of the yield curve);
(8) delays in taking management actions;
(9) catch limits not being fully utilised;
(10) unbalanced sex ratios in catches.

For some of the above factors, the output would be presented as for the initial screening procedure. For others, such as (8), it may not be necessary to model the process precisely, but rather give a general indication of the likely effect. Having passed such qualitatively more stringent tests, simulations of these procedures would be repeated using more trials (e.g. 2,000) to improve estimates of risk levels and tune management parameters. It was recognised that it was not possible at this Workshop to determine precisely the procedures to follow for second stage testing. The above guidelines will need to be modified in the light of results from the initial screening process.

## 7. RECOMMENDATIONS TO THE SCIENTIFIC COMMITTEE

The Workshop recommends that evaluation of possible management procedures be based on the guidelines developed under Item 6.4.

Given this uniform approach to the presentation of results, the question of standardised software was considered. It was agreed that, in the early stages of investigation, the benefits of having such software may be outweighed by a lack of efficiency due to unfamiliarity with a chosen standard language (e.g. FORTRAN) and technical problems in combining standard modules with the particular management procedure.

Nevertheless, the Workshop recognised the utility of using the IWC Secretariat to validate and hold programs. Accordingly the Workshop recommends that this be done for second stage testing. It also recommends that the Scientific Committee consider the implications of this in their discussion of the enhancement of the IWC computing facility and replacement of the senior analyst/programmer.

In this regard the Workshop recalled the experience with North Pacific sperm whale models and noted the value of:
(1) providing detailed instructions on the work to be carried out;
(2) monitoring progress with a small steering group.

It recommends a similar procedure be followed for management computing work.

Development of alternative management procedures, and their evaluation as outlined under Item 6.4, requires considerable computer time. During the initial stages it will be most efficient for the work to be carried out by the author. From a financial standpoint, three options are available for seeking assistance:
(1) submission of a research proposal following either the procedure outlined in Rep. int. Whal. Commn 35:34-5 or as a recommended proposal developed within the Committee;
(2) application for a discretionary grant;
(3) use of the Secretariat to provide the necessary output from the program supplied by the author.

With respect to (2), the Workshop recommends that there be a limit of $£ 1,000$ per project for initial screen testing of possible management procedures to be available as a grant for authors to use on their local systems.

The guidelines for initial screening procedures may need to be modified in the light of experience. This could be accelerated if the results of initial attempts to utilise the guidelines are available at the coming 1987 Scientific

Committee meeting. The Workshop welcomed the offer by Sakuramoto to try to provide some results for the Sakuramoto and Tanaka procedure at that meeting.

It was agreed that progress on the development of management procedures would need to be reviewed at a future Workshop. The timing of such a Workshop was conditional on three factors:
(1) progress made by individual authors;
(2) the deliberations of the Commission regarding the Comprehensive Assessment and revised management procedures;
(3) enhancement of the IWC computing facility.

With respect to (3) it is important that the member of the Secretariat who will carry out any second stage testing fully participate in any future workshop. If, as seems likely, this work will be carried out by a new appointee, then, allowing for a decision by the Commission to recruit such an appointee and the time taken to do this, it may not be possible for this Workshop to take place before the 1988 Scientific Committee meeting.

## Annex A <br> Participants

| ICELAND | K. Shima | SPAIN | IWC |
| :--- | :--- | :--- | :--- |
| T. Helgason | K. Yamamura | M. Quiroga | G.P. Donovan |
| K. Magnússon |  | S. Lens | R. Gambell |
| J. Sigurjónsson <br> G. Stefánsson | NETHERLANDS |  |  |
|  | K. Lankester |  |  |
| JAPAN |  | UK |  |
| H. Kishino | NORWAY | J.W. Horwood | INVITED PARTICIPANTS |
| T. Miyashita | N. Øien |  | D.S. Butterworth |
| T. Nakamura |  |  | J.G. Cooke |
| S. Ohsumi | SEYCHELLES | S.A. | W.K. de la Mare |
| K. Sakuramoto | S.J. Holt | M.F. Tillman | O.L.R. Jacobs |
|  |  |  | G.P. Kirkwood |

## Annex B Agenda

1. Convenor's opening remarks
2. Election of Chairman and rapporteurs
3. Adoption of Agenda
4. Arrangements for meeting
5. Review of documents
6. Consideration of scientific aspects of simulation and theoretical studies of alternative management strategies 6.1 Review of the new management procedure (NMP)
6.2 Alternative management strategies
6.2.1 de la Mare's studies
6.2.2 Cooke's approach
6.2.3 Sakuramoto and Tanaka's approach
6.2.4 Magnusson's approach
6.3 Management procedures and stochastic control theory
6.4 Evaluation of proposed management procedures
6.4.1 Initial screening procedure
6.4.2 Second stage testing
7. Recommendations to the Scientific Committee
8. Adoption of report

# Annex C List of Documents 

## 1. Background Documents

Cooke, J.G. 1986. The assessment and management of whale stocks. Paper SC/A86/CA9 presented to the IWC Scientific Committee Special Meeting on Planning for a Comprehensive Assessment of Whale Stocks.
De la Mare, W.K. 1986. Simulation studies on management procedures. Rep. int. Whal. Commn 36: 429-50.
De la Mare, W.K. 1986. Further simulation studies on management procedures. Paper SC/38/O 3 presented to the IWC Scientific Committee, June 1986.
Sakuramoto, K. and Tanaka, S. 1986. A simulation study on management of whale stocks considering feedback systems. Paper SC/38/O 10 presented to the IWC Scientific Committee, June 1986.
Tanaka, S. 1980. A theoretical consideration on the management of a stock-fishery system by catch quota and on its dynamical properties. Bull. Jap. Soc. Sci. Fish. 46: 1,477-82.
Tanaka, S. 1982. The management of a stock-fishery system by manipulating the catch quota based on the
difference between present and target stock levels. Bull. Jap. Soc. Sci. Fish. 48: 1,725-9.
Tanaka, S. 1986. On a practical method for stock management. Paper SC/A86/CA6 presented to the IWC Scientific Committee Special Meeting on Planning for a Comprehensive Assessment of Whale Stocks, April 1986.

## 2. Meeting Documents

## SC/M87/M

1 Sakuramoto, K. and Tanaka, S. Further simulation study on management of whale stocks considering feedback systems.
2 De la Mare, W.K. An overview of some recent work on the management of exploited whale populations
3 Lankester, K. The stock identities of whales: implementation in the Comprehensive Assessment
4 Magnusson, K. A note on a feedback strategy to regulate catches from a whale stock

# Annex D <br> Management Procedure from the Point of View of Stochastic Control Theory 

O. Jacobs

The problem of managing a population could be represented as the problem of controlling a system (Fig. 1) having input $C$ (the catch quota), output $F$ (the effort expended in obtaining the specified catch) and internal state $P$ (the population level). The system is subject to uncertainty and therefore should be controlled by feedback.

Experience suggests that two levels of feedback may be required; a basic-level feedback control subject to higher-level jacketing software (Fig. 2). Stochastic control theory offers general concepts which may help with the basic control, as discussed below. Jacketing software, for example near-extinction thresholds or use of supplementary information about population level, is usually problem-specific and best developed by local experts.

Stochastic control theory is based on a model of the controlled process consisting of two blocks (Fig. 3) relating input $u$, dynamic state $x$ and measured output $y$ :
(i) Dynamics $\quad x(t+1)=f_{1}\left(x, u, r_{1}\right)$
(ii) Measurement $\quad y=f_{2}\left(x, u, r_{2}\right)$
(iii) Some specification of desired performance is also necessary.

A general result is that a feedback controller must perform two separate functions (Fig. 4):
(a) Estimate the dynamic state, i.e. generate a conditional probability density $\mathrm{p}(\mathrm{x} \mid \mathrm{y})$ or some approximation thereof;
(b) Implement feedback control action $u(p(x \mid y))$ determined as a function of the estimator output $\mathrm{p}(\mathrm{x} \mid \mathrm{y})$.

Design of both (a) and (b) should be matched to the specifications (i), (ii) and (in the case of (b)) (iii) above. If the functions $f_{1}$ and $f_{2}$ are both linear and there is a quadratic cost function, it is well known that the estimator should be a Kalman filter, that the control law should be linear in the estimates and that certainty equivalent control (neglecting accuracy of estimates) is optimal.

In the application to population management the variables correspond as follows:
$u=C \quad$ The actuating variable is catch quota.
$\mathrm{x}=\mathrm{P} \quad$ The dynamic state is population level.
$y=F \quad$ The measurable output is effort to obtain $C$.

Typical functions might be:
(i) $\mathrm{P}(\mathrm{t}+1)=\alpha \mathrm{P}+\mathrm{R}\left(\mathrm{P}, \mathrm{r}_{1}\right)-\mathrm{C}$ where $0<\alpha<1$ accounts for natural mortality, the recruitment $\mathrm{R}=\mathrm{P}(1-\mathrm{P} / \mathrm{Pmax}) \mathrm{r}_{1}$
and $r_{1}$ might be lognormal.
(ii) $\mathrm{F}=\mathrm{r}_{2} \mathrm{C} / \mathrm{P}$ where $r_{2}$ might be lognormal.

These functions are both non-linear and it is quite unlikely that a control which would be suitable for a linear system would give the best regulation. It is also possible that, with information about the current state x as poor as it is in whaling, even the 'best' regulation may not be very good.

Based on the above considerations, I make the following proposals.
(1) Agree a simplest possible mathematical model capturing essential features of the controlled process, including dynamics $f_{1}$, measurement $f_{2}$, and random variables $\mathrm{r}_{1}, \mathrm{r}_{2}$.
(2) Investigate controller structures to match the non-linear controlled process.
(3) Investigate estimators to match the non-linear controlled process. What about recursive estimators having the predictor-corrector structure seen in the Kalman filter?
(4) Do not confuse jacketing software with the basic controller.


Figs. 1-4 (see text).

# Comprehensive Assessment Workshop on Management Procedures 

The Workshop was held at the Fisheries Laboratory, Lowestoft, UK, from 6-10 February 1989. A list of participants is given in Annex A. Horwood welcomed participants on behalf of the hosting laboratory.

## 1. CONVENOR'S OPENING REMARKS

Kirkwood outlined the background to the Workshop, stressing the importance of the topic under consideration to the Comprehensive Assessment (Rep. int. Whal. Commn 39:131-47).

## 2. ELECTION OF CHAIRMAN AND RAPPORTEURS

Kirkwood was elected Chairman. Allison, Butterworth and de la Mare were appointed rapporteurs. Donovan assisted in the final stages of report preparation.

## 3. ADOPTION OF AGENDA

The Agenda adopted is given as Annex B. It was clarified that Item 10 concerned consideration of the best ways in which proposals might be presented to the Commission.

## 4. ARRANGEMENTS FOR MEETING

The Workshop agreed to a work schedule suggested by the Chairman. He also stressed the need for participants to remember that English was not the primary language of a number of participants.

## 5. REVIEW OF DOCUMENTS

Annex C lists papers presented to the meeting.

## 6. FIRST STAGE SCREENING OF ALTERNATIVE MANAGEMENT PROCEDURES

### 6.1 Procedures presented

Five documents were presented to the meeting describing progress and giving partial or complete first stage screening results for suggested management procedures: Punt-Butterworth (SC/F89/M2); Sakuramoto-Tanaka (SC/F89/M4); de la Mare (SC/F89/M5); Cooke (SC/F89/M6); and Magnusson-Stefansson (SC/F89/M9). Detailed summaries and comments upon these procedures prepared by their authors are given in Annex D.

Certain features are common to some or all of the procedures. All use a catch limit management system. Three are based on fitting population models to relative and/or absolute abundance data (Punt-Butterworth, de la Mare, Cooke), whereas the other two (SakuramotoTanaka and Magnusson-Stefansson) use cmpirical approaches involving levels and/or rates of changes of a relative abundance index.
The Cooke and de la Mare approaches require only absolute abundance data, such as sightings survey estimates, although these data are often used only as a
relative index. The Sakuramoto-Tanaka and Magnusson-Stefansson procedures require only a single absolute abundance estimate, followed by a time series of relative abundance indices which could be, for example, CPUE (catch per unit effort) or sightings survey data. Only the Punt-Butterworth procedure uses both CPUE and sightings survey data on a continuing basis.

Most of the procedures need to set somewhat arbitrary catch limits over an initial period (of the order of a decade in duration), until sufficient data are available for their primary control laws to be implemented. The population models used by the Punt-Butterworth and de la Mare procedures are delay difference equations of the same form as the model used to generate data for the simulation trials, whereas the Cooke procedure uses a simpler Schaefer model; these three procedures fix a population level (relative to the carrying capacity level, K ) below which the population is protected, but the levels chosen differ. The other two (empirically based) procedures do not incorporate the concept of a protection level. All procedures except that of Magnusson-Stefansson include specific additional restrictions intended to smooth catch limit variations. In the de la Mare and Cooke procedures, catch limits are progressively reduced in a manner explicitly related to the precision of abundance or population level estimates.

The Punt-Butterworth procedure uses a variant of the control law of the New Management Procedure (see Section 7), but attempts to reduce catch limit fluctuations by expanding the range of population size over which catches change from zero to a maximum. A population model is used to estimate MSY (the maximum sustainable yield), but if this cannot be achieved successfully, a pre-fixed value of MSY\% (the ratio of the MSY to the population size at which MSY is achieved, i.e. MSYL, expressed as a percentage) is chosen.

Catch limits in the Sakuramoto-Tanaka procedure are changed depending on the difference between an index of population size and a target level, and also the rate of change of this index with time. The procedure incorporates an algorithm to move the originally chosen target level progressively towards MSYL.

The de la Mare procedure calculates catch limits using estimates of current population size relative to K and of replacement yield (RY). Catch limits are set above or below RY so as to move the population towards a target level. The target level is set higher for less precise model estimates of current population size relative to $K$.

Particular features of Cooke's procedure are the simple form of the model and estimation process used, and the incorporation of a prior distribution for the specific growth rate parameter $r$ in this process, in order to reduce the variability of estimates when there are few data. Catch limits depend on estimates of stock depletion according to a control law.

The Magnusson-Stefansson procedure adjusts catch limits according to the rate of change of an index of population size with time in a manner that assures
stabilisation of the population at a non-zero level. After certain periods the catch limit is changed by a fixed percentage so that the stock approaches a new equilibrium, in such a way that the population is moved over time towards MSYL.

The results for the first full set of first stage screening simulation trials specified at the 1987 Reykjavik Workshop (Rep. int. Whal. Commn 38:163-70) were presented for four of the procedures; only a partial set was presented for the Magnusson-Stefansson procedure. It was agreed that the four procedures that had finished first stage screening merited continued consideration and investigation under second stage screening; first stage screening for the fifth procedure should be completed.

In the general discussion it was observed that, other things being equal, simpler procedures could be more easily explained to the Commission and others concerned and eventually written into the Schedule.

Specific suggestions made for further analyses of two of the procedures were: (1) that the behaviour of the Punt-Butterworth procedure be investigated when CPUE data were not incorporated and when their 'bottom line' estimator (see Annex D1) was used at all times, and (2) that the prior distribution for the specific growth rate parameter $r$ in the Cooke procedure be modified to attempt to improve the reaction time of the procedure.

It was noted that all five procedures envisaged management based on catch limits and that none had considered the possibility of regulation by effort limitation. Reasons advanced for the unsatisfactory nature of effort limitation as a basis for management were: (1) that indices of effort most likely to relate to fishing mortality were derived (e.g. searching time) rather than raw statistics, which could not form the basis of a limitation to be imposed in a practical situation, and (2) the possible non-linearity in the fishing mortality-effort relation, which would lead to catches not falling in proportion to population size and hence to a greater population reduction than might be intended.

Shepherd presented Annex E1, in which he suggested that it would be desirable to test a 'Constant F' control law (although the value of F would be updated as appropriate estimation became feasible) and a 'Cautious' control law for which catches varied quadratically with population size. A variant of Cooke's estimation procedure could be used to provide the estimates required. The principal objectives of these control laws were simplicity and a smoother time series of catches. In a further document (Annex E2), concern was expressed that control laws which switched quickly between high and low (or zero) catches could lead to instability, and it was suggested that this matter warranted further study.

Reservations expressed by some participants concerning these suggestions were that smooth catch trends might preclude the probing necessary for precise estimation of surplus production potential and hence lead to under-utilisation of a stock, and that the instabilities evident in the performances of certain procedures were not primarily (if at all) consequences of deterministic dynamic effects. They also noted that there was no indication that the instabilities outlined in Annex E2 existed in the procedures examined. The Workshop agreed that the authors of the five proposed procedures should take note of the suggestions underlying the control laws suggested in Annex E1 and the associated rationale, when considering further modifications to their procedures.

### 6.2 Finalisation of first stage screening trials

The authors of all but the Magnusson-Stefansson procedure reported that they had essentially completed first stage screening, although minor improvements to their procedures under such screening could still be attempted.

However, in consideration of the results of these trials it became clear that certain ambiguities in the specification of the trials (Rep. int. Whal. Commn 38:165-7) would give rise to problems in making a proper comparison of the results. Such problems were mainly associated with the lack of clarity regarding specification of error sizes and structures of a procedure that used more than one set of abundance indices. It was agreed that a final set of first stage screening trials should be conducted which would utilise one or both of two sets of abundance indices to be referred to as:
(i) 'CPUE' - available annually (provided a catch was taken) with a log-normal error structure and coefficient of variation as described in Annex F;
(ii) 'Sightings survey' - only available every fifth year (may be taken to provide either an absolute or a relative index of abundance) with a log-normal error structure and with all trials to be conducted both for $\mathrm{CV}=0.2$ and $C V=0.4$.

The error structure for the CPUE data is a simpler variant of the gamma distribution used earlier, and shows the effect of increasing variance as catch sizes become small. The reason for no longer specifying trials with sighting surveys performed every year is that surveys with CVs of 0.2 every fifth year will be nearly equivalent to ones with CVs of 0.4 every year. It was suspected that procedures using sightings surveys might experience problems for $\mathrm{CV}=0.4$; thus trials were to be repeated with the smaller $\mathrm{CV}=0.2$ as the results might have important implications for minimum intensity or frequency requirements for such surveys.

Trials for 'rehabilitation' (i.e. initially depleted stock) cases ( $\mathrm{P}_{0}=0.3 \mathrm{~K}$ ) need only consider the situation where no historic CPUE data are available. This is because (possibly) acceptable historic CPUE series of such a kind are only available for very few real whale stocks, and procedures must in any case be able to be applied in their absence. In addition existing results showed the performance of some procedures to be relatively insensitive to the availability or otherwise of historic CPUE data. Information on historic catch, as opposed to CPUE, data remains available. Further investigation of cases with MSYL $=0.8 \mathrm{~K}$ will be deferred to second stage screening (see Section 8).

Thus the 14 case protocol agreed at the 1987 Reykjavik Workshop is reduced to examination of the eight cases shown in Table 1.

Table 1

| Trial no. | $\mathrm{P}_{0}$ | MSY\% | MSYL\% | Sightings <br> survey CV |
| :---: | ---: | :---: | :--- | :---: |
| 1 | K | $1 \%$ | 0.6 K | 0.4 |
| 2 | 0.3 K | $1 \%$ | 0.6 K | 0.4 |
| 3 | K | $4 \%$ | 0.6 K | 0.4 |
| 4 | 0.3 K | $4 \%$ | 0.6 K | 0.4 |
|  |  |  |  |  |
| 5 | 0.3 K | $1 \%$ | 0.6 K | 0.2 |
| 6 | K | $1 \%$ | 0.6 K | 0.2 |
| 7 | 0.3 K | $4 \%$ | 0.6 K | 0.2 |
| 8 |  | $4 \%$ | 0.6 K | 0.2 |

Essentially then, four 'base cases' with $\mathrm{P}_{\mathbf{0}}=\mathrm{K}$ or 0.3 K and MSY\% $=1 \%$ or $4 \%$ (and all with MSYL $=0.6 \mathrm{~K}$ ) are to be examined for different degrees of precision of sightings surveys.

For procedures using CPUE data as well as one or more sightings surveys, these eight cases must be examined for underlying situations for which CPUE is proportional to abundance and for which CPUE is proportional to the square root of abundance, giving a total of 16 cases in all. For procedures using CPUE data only, these two relationships with abundance must still be investigated, but since the two possibilities for sightings survey CVs fall away, only eight cases remain to be examined. In all other respects, specifications for the trials are as laid out in Rep. int. Whal. Commn 38: 165-7.

The following information may be assumed to be known to a management procedure:
(i) the series of historic catches (i.e. for those cases with $\mathrm{P}_{0}=0.3 \mathrm{~K}$ );
(ii) that the population was at carrying capacity level at the start of exploitation (i.e. $\mathrm{P}_{0}=\mathrm{K}$ or $\mathrm{P}_{-30}=\mathrm{K}$ as appropriate).
Further information that may be assumed known, such as the age at maturity, is detailed in the Report of the 1987 Workshop. A manager may not know a priori whether the sightings survey CV is 0.2 or 0.4 , nor the form of the relationship between CPUE and abundance.

It was recognised that while the large amount of data output and the 15 histograms required to be produced for each trial as agreed at the 1987 Workshop had assisted in the development of procedures thus far, it was essential that this volume of statistics be substantially reduced in presenting the results of future trials, to facilitate comparisons. It was agreed that the following eight statistics were the most important and would need to be reported for the 100 stochastic simulations for each trial:
(1) and (2) average and standard deviation of total catch over first 100 years of management;
(3) average root mean square inter-annual catch variability over the first 100 years of management;
(4) average of the CV of the catch for the first 100 years of management over the 100 simulations;
(5) and (6) average and standard deviation of final population size after 100 years of management;
(7) and (8) average and standard deviation of the lowest population size over the first 100 years of management.

Only one graph is required, showing population and catch for deterministic simulations, for each of the four base cases (or eight cases if CPUE data are used) over the first 100 years of management. The reason for concentrating on results over a 100 year time scale is that the slow dynamics of whale populations mean that certain important aspects of population response to a management procedure are not necessarily apparent over shorter periods. It was noted that the plots of deterministic simulations would give an indication of the behaviour of the procedure over the earlier part of this period.

The list above constitutes the minimum output requirement; naturally the authors of different proposed procedures are free to provide any additional output that they may consider desirable.

In order to assist further in the comparison of procedures which imposed a restriction on the maximum inter-annual catch fluctuation (Punt-Butterworth, Sakuramoto-Tanaka
and de la Mare), it was agreed that the associated authors would all use a maximum allowed variation of $\pm 20 \%$ of catches from year to year (except that such a restriction would not apply to any annual catches that did not exceed 20 whales).

Thus if the basic control law provides an initial catch limit $C^{\prime}{ }_{n}$ for year $n$, and $C_{n}^{\prime}>20$, then:

$$
\begin{aligned}
& 0.8 C_{n-1}>C_{n}^{\prime} \Rightarrow C_{n}=0.8 C_{n-1} \\
& 1.2 C_{n-1}<C_{n}^{\prime} \Rightarrow C_{n}=1.2 C_{n-1}
\end{aligned}
$$

This would not preclude additional rules related to such variability in certain circumstances.

The authors of all five procedures indicated that they intended to report the results of these final first stage screening trials at the 1989 Scientific Committee meeting. It was agreed that modification of procedures before that time could be attempted.

### 6.3 Basis for comparison of procedures

SC/F89/M3 suggested that consideration needed to be given to systematic methods for comparing the performance of different proposed management procedures, because of the large number of attributes which had been indicated as pertinent to the measurement of such performance (even given the reduction in statistics required for reporting agreed above). Attempts to undertake this in terms of a multi-attribute utility function were not seen as appropriate by the authors of SC/F89/M3 because the use of weighted sums of attribute values can lead to severely suboptimal solutions, and correct evaluation of such a utility function is an enormous if not impossible exercise. Given that research shows there are severe constraints on the number of attributes in terms of which different procedures can be consistently compared, specification of a much smaller set of the more important attributes becomes important as does the use of a systematic comparison process, although no such method can entirely remove subjective aspects from the process of selection between procedures.

The authors suggested that an algorithm which may be appropriate in this situation (comparing a small number of procedures in terms of a large number of attributes) is that developed by Roy and co-workers (e.g. Roy and Vincke, 1981, European J. Op. Res. 8: 207-18). The idea of the algorithm is to select the procedure which is better than the others in terms of many of the attributes, but is not disastrous for any of them. The respective measures of these considerations are provided by a Concordance and a Discordance matrix, and a specific comparison process is defined. This enables successive procedures to be eliminated, although it may not necessarily provide a single 'best' procedure. An example of application of the algorithm to select between variants of the Punt-Butterworth management procedure was provided. The intent of presentation of the paper at the Workshop was not that a comparison using the algorithm suggested should be attempted immediately, but rather that the attention of the Scientific Committee should be drawn to the comparison problem, and that they consider seeking advice from specialists in the field of multi-criteria optimisation and decision support systems.

SC/F89/M8 addressed the question of formulating the problem of finding an optimal management strategy in terms of optimal control theory. Standard optimal control theory involves defining a simple optimality criterion
which is to be maximised. Because of the problem of balancing short term catch gains against longer term detriments to the stock, and in particular, the extreme dependence of the result on the value chosen for the discount rate, Cooke concluded that this is not an appropriate formulation of the problem. As an alternative he suggested that some maximum acceptable risk of stock depletion be defined as a constraint subject to which an optimality criterion including catch is maximised. Horwood considered that, provided a few essential criteria for management could be specified, then near-optimal techniques might be applicable to the problems of whale management, which could incorporate risk of depletion and costs of change in quotas and would perform better than present models. Horwood also introduced a paper by Ballance, Jacobs and Horwood (1988, State estimation in regulating a harvested population, pp. 839-44 In: 8th IFAC/IFORS Symposium on Identification and System Parameter Estimation, Beijing.) which developed a one-step optimal control law for fisheries by seeking to balance variation in catch and effort.

The Workshop concluded that the development of an evaluation procedure for comparing the performance of various potential management procedures is very important, and that this should proceed in parallel with second stage screening. Selection among different management procedures proposed was not appropriate at this stage as the intent of first stage screening had only been to determine whether a proposed procedure performed adequately so that its progression to second stage screening was justified. In fact it might be the case that no single one of the present proposed procedures might be recommended in due course, but rather some combination selecting the best features from each. Future comparisons would also need to take both long and short term performance measures into account. The Workshop further noted that a reduction in the number of cases to be simulated along with even greater reductions in the number of statistics used to describe the results (as had been attempted in the revised specifications above) would make comparative evaluation less difficult.

## 7. STRENGTHS AND WEAKNESSES OF NMP AND ALTERNATIVE MANAGEMENT PROCEDURES

The Workshop decided that a set of first stage screening results for a management procedure based on the New Management Procedure (NMP) would represent a useful baseline against which to compare potential revised management procedures. Because the Scientific Committee found it problematic to apply the NMP, it is difficult to decide exactly how to simulate the way in which the procedure should be applied in the circumstances assumed by the other management procedures in the screening tests. In addition, the NMP does not specify how to set catch limits in the absence of an acceptable estimate of MSY. The Workshop agreed that the interpretation of the NMP used in simulation studies by de la Mare (Rep. int. Whal. Commn 37:429-50) should form the basis for the screening tests. In those studies, MSY and depletion were estimated by fitting a population model to time series of either absolute or relative abundance data. Catch limits are set at $90 \%$ of the MSY estimate or at $5 \%$ of the initial abundance estimate, whichever is the less. However, if the MSY estimate would lead to a catch limit double the average catch over the preceding ten years, then the latter
is used as the catch limit. Initial catch limits in a previously unexploited stock are set at $90 \%$ of $4 \%$ of MSYL. The Workshop stressed that if the trials of the NMP-like procedure appeared to be reasonably successful this would not mean that the NMP is satisfactory as it is now specified in the Schedule, or in the ways in which implementation has been attempted in the past.

As discussed in Section 6, the Workshop agreed that it was not appropriate to examine the relative strengths and weaknesses of the five proposed procedures at this stage.

## 8. DEVELOPMENT OF SECOND STAGE SCREENING PROTOCOLS

The aims of screening procedures are twofold. The first aim is to test potential management procedures against a range of problems arising from failures in assumptions about stock dynamics and data, in order to determine whether they adequately attain management objectives. Sources of uncertainty which lead to serious failures in potential management procedures require either that the procedure be modified to cope with the uncertainty, or that other research be designed for resolving the uncertainty.

The second aim is to produce measures describing the performance of potential management procedures so that the adoption of a procedure or elements of a procedure can be based on objective criteria.

As a general principle it was considered that the problems management procedures face that are of greatest concern are those representing uncertainties that are difficult to resolve. A simple example is the uncertainty about the true relationship between stock size and sustainable yield. The Workshop decided that such problems should be given first priority in second stage screening. Problems which could be solved using information which could be reliably obtained, at least in principle, had a lower priority.

The Workshop adopted the principle that screening tests should be difficult for potential management procedures to handle, but not obviously impossible. This means that not all tests are necessarily designed to be fully plausible or realistic. This philosophy is based on the idea that if a procedure can cope with a severe form of a problem then it is usually unnecessary to be concerned with such problems in less severe forms. If a procedure fails with a severe form of a problem, then it becomes necessary to explore the degree to which the problem might be encountered in practice. From there it must be decided whether to modify the procedure or to conclude that there are no cases where the problem could arise with the degree of severity which would lead to the breakdown of a management procedure.

It was noted that the management procedures examined so far were directed towards the management of baleen whale stocks. The Workshop agreed that some of the problems which could arise in the development of management procedures for odontocetes required special attention, but decided that this could be delayed until the baleen whale procedures are well advanced in second stage testing.

The Workshop adapted the table from SC/F89/M7 as a summary of the types of problems and uncertainties which have either arisen in past whale assessments or which may lead to failure in the assumptions typically employed in various types of assessment procedures (Annex G). Of these problems, those that are both potentially important but require only minimal modifications to the existing
computer programs were selected for examination in the first phase of second stage screening. These are generally problems affecting the properties of the estimation procedures used in the stock assessments.

The Workshop recognised that the design of the second stage screening procedure must be an iterative process. For the first phase of this process the following tests were devised. In all cases, the four base cases described in Section 6 are to be tested, with the CV on the sighting surveys set at 0.2 and the CPUE data being proportional to the square root of stock abundance. Year 0 refers to the start of the true population simulation for all four base cases. It was suggested that the screening tests might be carried out more efficiently using partial factorial designs.

Problems which arise from incorrect stock identification and concentration of whaling within a small part of a management area have been recognised as being of great potential difficulty for management procedures (e.g. Rep. int. Whal. Commn 37:147-57).

The Workshop identified two simple tests ( 5 and 6 below) as the first step in examining problems which arise from stock misidentification.

## (1) Incorrectly specified population model

Test $1 a$
Left peaked yield curve using Pella-Tomlinson model with MSYL $=0.4 \mathrm{~K}(\mathrm{z}=0.0188)$.

Test $1 b$
Right hand peaked yield curve using Pella-Tomlinson model with MSYL $=0.8 \mathrm{~K}(\mathrm{z}=11.22)$.

## (2) Time variable parameters

Carrying capacity to vary sinusoidally with a period of 100 years, a minimum of 5,000 and a maximum of 15,000 .

Test $2 a$
Cycle minimum occurs in year 0 .
Test $2 b$
Cycle maximum occurs in year 0 .
(3) Bias in absolute abundance data

Test $3 a$
Multiplicative bias of 1.5 .
Test $3 b$
Multiplicative bias of 0.5 .

## (4) Trends in catchability for CPUE data

Test $4 a$
Catchability coefficient increases linearly from initial value to twice initial value in year 99 .

## Test $4 b$

Catchability coefficient declines linearly from initial value to half initial value in year 99 .

## (5) Two separate stocks managed as a single stock

A single putative stock is made up of two initially equal isolated stocks with the parameters specified in the initial screening protocol (i.e. the two stocks have the same values of K, MSY\% and other biological parameters). Estimates of absolute abundance come from surveys which cover the full range of both stocks. Only one of the two real stocks is exploited.

Test $5 a$
Both stocks are at carrying capacity K.
Test $5 b$
Exploited stock of 0.3 K , unexploited stock at K .

## (6) A single stock managed as two separate stocks

The true stock is divided into two putative stocks each containing half the population. One of the two putative stocks is exploited and catches are taken so that the true stock is reduced to 0.65 K . This means that if the two putative stocks were really isolated, the exploited one would have been reduced to 0.3 K . MSY rates of $1 \%$ and $4 \%$ are used for the true population, and perfect mixing occurs between the two halves of the stock.

More detailed tests involving incorrect stock identification and ecological interactions require more complex multistock mixing and competition models. However choice of appropriate parameter values for such models is better carried out when participants have had the benefit of experience in testing the simpler models above. Accordingly, the Workshop was unable to specify further tests on this class of problems. The Workshop recommends that the developers of procedures begin to construct tests along the lines proposed in SC/F89/M7 for future presentation, so as to guide the specification of tests suitable for inclusion in the second stage testing procedures.

## 9. WORK PLAN FOR IMPLEMENTING SECOND STAGE SCREENING - INCLUDING BUDGET PROPOSALS

The Workshop recommends that the revised first stage screening be carried out on the five procedures so that complete and comparable results are available at the 1989 meeting of the Scientific Committee. Since the procedures may require slight modifications in the light of these tests, it will be most efficient for the work to be carried out by the individual authors. Authors are also encouraged to begin some of the second stage testing if possible, and to develop suggestions for further second stage screening. Time will be needed at the forthcoming 1989 Scientific Committee meeting to present and review the revised first stage testing results and any others available, and to plan and specify further second stage tests. Running of the proposed simulation tests requires a considerable amount of computer time, and the Workshop recommends that authors be able to apply for a discretionary grant to cover costs up to a limit of $£ 1,000$ per project.

It will be necessary to hold a Workshop early in 1990 to review the results of the next phase of second stage screening, design its final phase and prepare for the report required as part of the Comprehensive Assessment. The budget required for the Workshop is estimated to be £15,000.

It had originally been envisaged that the Secretariat would perform the second stage testing, but since the programs are not yet in final form it was agreed that the initial second stage testing would be best carried out by the authors. The Workshop recommends that the Secretariat be responsible for running first stage screening tests on an implementation of the NMP (see Section 7) to provide comparison with the new procedures. Allison should also develop a common control program into which the various management modules can be fitted. This will ensure that
exactly the same information is available to each management module and will facilitate implementation of different scenarios in further second stage testing. In preparation for this it was suggested that it would be useful if the control program be written in a way which anticipates multi-stock analysis. The interface between the calling control program and the management modules was agreed, as was a standard version of a random number generating routine in order that the program be implemented consistently on all different machines. Allison undertook to send out copies of this program as soon as possible. The authors would be responsible for altering their programs into this format.

Recent versions of computer programs for four of the five management procedures are currently held by the Secretariat. The 1987 Workshop envisaged validation of the procedures by the Secretariat at this stage, followed by second stage screening. However, the revised strategy for second stage screening developed at this Workshop allows for further modification of the procedures by their authors. The Workshop agreed that it is more sensible for validation of the programs by the Secretariat to take place when the first phase of second stage screening has been completed, i.e. when the procedures are nearer their final form. This matter, and the details of the validation itself, should be addressed at the 1989 Annual Meeting.

## 10. PRESENTING PROPOSALS TO THE COMMISSION

The Workshop noted that half a day at the Annual Meeting had been allocated for discussions of the Comprehensive Assessment by a Joint Scientific and Technical Committee Working Group, and considered what form of presentation to the Group regarding progress in the development of revised management procedures might be desirable (Rep. int. Whal. Commn 39:14-15). It was agreed that separate detailed presentations by an author of each of the five procedures under present consideration would take too long and would not be appropriate. Nevertheless, even though development work is by no means complete, it was seen as essential that some form of general report of what work is in progress should be made. Furthermore, in view of the importance of the matter, the Workshop suggested that consideration should be given to making a short presentation directly to the Commission, possibly as a part of the report of the Chairman of the Scientific Committee at the Commission's meeting.

It was suggested that such a presentation should be kept at a relatively simple level and should concentrate on the philosophy underlying the procedures being developed. Possible aspects to emphasise might be: the trade-offs between total catches, the risk of depleting the population to a low level and the extent of variability of catches from one year to the next; the sorts of data required to implement the new procedures under investigation; and the fact that work on these procedures was progressing in a convergent manner. It was agreed that it would premature to give any detailed results of current trials of procedures in such a presentation, as all procedures were likely to be modified in the light of results obtained from second stage screening trials.
More specific inputs from the Commission on policy aspects would be needed in due course to facilitate the finalisation of proposed procedures. While it is
inappropriate to foreclose on possible options by asking specific questions of the Commission at this time, it is important that the presentation provides the Commission with some idea of the sorts of questions to which answers will be required in the future. Advice should also be given that the process of development and evaluation of procedures was more complex than had been envisaged at the 1987 Workshop.

The Workshop considered that Kirkwood was clearly the most appropriate person to make such a presentation. Kirkwood agreed to do so provided he was able to attend the meeting.

## 11. RECOMMENDATIONS TO THE SCIENTIFIC COMMITTEE

The Workshop agreed to the recommendations set out below.
(1) The revised first stage screening should be carried out for the proposed procedures and the results presented to the 1989 Scientific Committee meeting.
(2) The Secretariat should carry out the work on first stage screening of the 'NMP' (see Section 7) and develop a common control program for the various management modules (see Section 9).
(3) A further Workshop should be held early in 1990 to review available results of second stage screening trials (see Section 9).
(4) Funds should be set aside for the Workshop proposed and to assist authors of procedures with computing costs (see Section 9).
(5) Time should be set aside at the 1989 Scientific Committee meeting to:
(a) review the results of finalised first stage screening trials for each proposed procedure (see Section 6.2);
(b) consider the question of comparing proposed procedures (see Section 6.3);
(c) consider results of initial second stage screening trials that might be presented, and determine the details of further such trials (see Section 8);
(d) consider the questions involving the validation of the programs of the proposed procedures (see Section 9); and
(e) make preparations for the 1990 Workshop (see Section 9).
(6) A short presentation on progress in the development of revised management procedures should be made to the Commission (see Section 10).
Finally, the Workshop strongly urges that Kirkwood should, if at all possible, continue as Chairman of future meetings (including the sub-committee meeting at the 1989 Annual Meeting) regarding the development of management procedures. This is because of the necessity for continuity of leadership and Kirkwood's unique position of possessing detailed knowledge of the subject but without being directly involved in the development of any of the procedures under consideration.

## 12. ADOPTION OF REPORT

The report of the meeting was adopted by the Workshop. It agreed that minor editorial amendments and clarifications, and finalisation of Annexes, could be
undertaken by Donovan after the meeting, in consultation with the Chairman and rapporteurs.

Kirkwood thanked the UK Commissioner for the invitation to host the meeting, and the Director and staff of
the Lowestoft laboratory for their hospitality. He also thanked the rapporteurs for their work.
The Workshop expressed its appreciation to Kirkwood for his customary efficient chairmanship.

## Annex A <br> Participants

| Iceland | Netherlands | J.G. Shepherd | Invited participants <br> K.G. Magnusson |
| :--- | :--- | :--- | :--- |
|  | K. Lankester | K. Stokes | W.K. de la Mare |
| Japan | Norway | USA | G.P. Kirkwood |
| H. Hiroyama | N. Øien | M.F. Tillman | A.E. Punt |
| T. Kasuya | T. Schweder |  | K. Sakuramoto |
| H. Kishino |  | IUCN |  |
| S. Misaki (Interpreter) <br> M. Morimoto | Seychelles | J.G. Cooke |  |
| T. Nakamura |  |  |  |
| S. Tanaka |  | IWC |  |
| K. Yamamura | J.W. Horwood | C. Allison |  |

## Annex B

Agenda

1. Convenor's opening remarks
2. Election of Chairman and Rapporteurs
3. Adoption of agenda
4. Arrangements for meeting
5. Review of documents
6. First stage screening of alternative management procedures
7. Strengths and weaknesses of NMP and alternative management procedures
8. Development of second stage screening protocols
9. Work plan for implementing second stage screening including budget proposals
10. Presenting proposals to the Commission
11. Recommendations to Scientific Committee
12. Adoption of report

## Annex C List of Documents

## SC/F89/M

1. COOKE, J.G. A note on computer programs for the simulation of whale stock management procedures.
2. PUNT, A.E. and BUTTERWORTH, D.S. A proposed whale stock management procedure.
3. STEWART, T.J., BUTTERWORTH, D.S. and PUNT, A.E. On comparing the performance of different proposed whale stock management procedures.
4. SAKURAMOTO, K. and TANAKA, S. Initial screening for an alternative management procedure of whale stocks.
5. DE LA MARE, W.K. Simulation studies on elements of a revised whale management procedure (SC/40/O 19).
6. COOKE, J.G. Simulation trials of a whale stock management procedure.
7. DE LA MARE, W.K. A range of factors for consideration in the design of a second stage screening protocol for revised whale management procedures.
8. COOKE, J.G. A note on the specification of the whale management problem.
9. MAGNUSSON, K.G. and STEFANSSON, G. A feedback strategy to regulate catches from a whale stock - a modified version and some simulations.

# Annex D <br> Author's Accounts of their Proposed Management Procedures 

## Annex D1. THE PUNT-BUTTERWORTH PROCEDURE

The harvesting algorithm component of this procedure is similar to that of the New Management Procedure (see Fig. 1), the fundamental difference being that the range of population levels over which catch limits are changed from zero to 0.9 MSY is widened from $(0.54 \mathrm{~K}, 0.60 \mathrm{~K})$ to $(0.2 \mathrm{~K}$, 0.7 K ), where K is the estimated unexploited population level. This widening is intended to promote greater inter-annual catch level stability in the presence of estimates of population size that fluctuate substantially from year to year because of the considerable noise level in the data. To enhance this catch level stability further, catches are not allowed to change by more than $25 \%$ from one year to the next, although greater decrements are possible if the population level is assessed to be less than 0.35 K .


Fig. 1. Diagrammatic illustration of the catch limit algorithm. The solid curve is the sustainable yield as a function of population size (Pella-Tomlinson form with MSYL=0.6K), and the dashed line indicates the catch limit corresponding to various population sizes. $\mathrm{Y}_{\mathrm{opt}}$ is the deterministic equilibrium point for this algorithm.

The associated estimation procedure, which provides values of the parameters and variables required to implement the harvesting algorithm, uses annual CPUE data and an unbiased absolute abundance estimate which is available in the first and every fifth following year. These data are fitted by a population model of the same form as the operating model (which provides the data for the simulation tests), which is cast in the form of a linear process error estimator to avoid unrealistically long computer time requirements to conduct the trials required. However, the fit to the model is rejected if the parameter values estimated fail to meet certain constraints, including ones which are biologically motivated such as a restriction that $0<\mathrm{MSY} \%<5 \%$ (assuming that MSYL $=0.6 \mathrm{~K}$ ). A hierarchy of simpler estimation procedures is then
attempted, culminating in a 'bottom line' observation error estimator, which fixes MSY\% $=1 \%$ in the fitting process and uses only absolute abundance estimates; the resultant population level estimates are then substituted into the harvesting algorithm with MSY\% fixed at $1.5 \%$ (see Fig. 2) to ascertain catch limits.


Fig. 2. Diagrammatic illustration of the original catch limit algorithm used with the 'bottom line' estimator, and the variant thereof. The solid curves are the sustainable various values of MSY\%. The dotted line indicates the 'bottom line' catch limit corresponding to various population sizes.

Results of the tests specified show relative insensitivity to the true MSYL, possible non-linearity in the CPUE-abundance relation and changes in its error distribution, and whether or not CPUE data are available prior to the implementation of the management procedure. Obtaining absolute abundance estimates every three years instead of every five makes no substantial impact on the performance of the procedure. Use of such estimates with negative biases of $30 \%$ leads to total catch reductions which are generally somewhat greater than $30 \%$. Increasing the population level for zero basic catch limit (see Fig. 1) from 0.2 K to 0.25 K results in little change in performance. Empirical estimation of the exponent in the CPUE-abundance relationship is successful, but as it is only viable after 30 years, little improvement in lowest population sizes (which generally occur before that time) is achieved.

The procedure performs satisfactorily for rehabilitation cases both for an MSY\% of $1 \%$ and of $4 \%$, and for development cases with an MSY\% of $1 \%$; however, the development case with an MSY\% of $4 \%$ is substantially under-utilised compared to the deterministic optimum
because the procedure produces inadequate contrast in the data to allow MSY\% to be well estimated. Considerable improvement in this respect is achieved by a variant of the bottom line estimator (see Fig. 2), which increases MSY\% for population levels greater than 0.7 K , though there is an accompanying large increment in the catch limit variability over the earlier years of exploitation.

Promising features of the procedure are the performance of the 'bottom line' estimator which prevents large stock reductions occurring while not precluding subsequent successful estimation of MSY\% to allow reasonable utilisation of resource production potential. In particular, the variant to this estimator which increases MSY\% for population estimates above 0.7 K , is successful in achieving a substantial improvement in resource utilisation for a scenario where many other procedures are unable to do this.

Disappointing aspects of the procedure are the poor performance of the process error estimator in estimating MSY\% (in tests comparing this to the ability of the corresponding observation error estimator), although this may be improved by increasing the extent of quadratic pre-smoothing of the CPUE data for the process error
estimator. A number of the safeguards against imprecise estimates that are built into the procedure are incorporated into the constraints applied to the fit from the process error estimator, do not have ready analogs for the observation error estimators more likely to be used in real situations. The overall estimation procedure is very complicated and so could not be simply explained; a condensation of this procedure, which might well result in rather little degradation in overall performance, merits investigation.

In assessing the performance of the procedure (and in partially tuning some of the control parameters), the attributes to which most attention was given were average final population sizes over a 100 year period, average total catches over this period, average lowest population sizes over this period, and catch variability over the first 10 years in the eight basic cases MSYL $=0.6 \mathrm{~K}, \mathrm{CPUE}$ error $\log$ normal $\mathrm{CV}=0.4$, initial population K and 0.3 K , MSY\% of $1 \%$ and $4 \%$, and CPUE proportional to population size and the square root of population size. (Little performance difference is evident whether or not historic CPUE data are available in the rehabilitation cases, so that attention can be focussed only on situations where they are not.)

## Annex D2. THE SAKURAMOTO-TANAKA (S-T) PROCEDURE

The S-T procedure is an entirely different system of whale stock management from the New Management Procedure (NMP); it manages whale stocks empirically. It does not require estimates of MSY\%, MSYL and RY, the components of the NMP that cause difficulties in its application to whale stock management. By comparison, the S-T procedure uses only a relative index of stock abundance for its adjustment of the catch. In other words, the catch limit is increased or decreased depending on whether the relative index is higher or lower than a pre-determined target level, and according to the increase or decrease of the relative index itself.

The formula below expresses this system:

$$
\frac{\mathrm{Y}_{\mathrm{t}-l}-\mathrm{Y}_{\mathrm{t}}}{\mathrm{Y}_{\mathrm{t}}}=\mathrm{h}^{*} \mathrm{~L}_{\mathrm{t}-l}+\mathrm{g}^{*} \mathrm{~K}_{\mathrm{t}-l} \quad \mathrm{t} \geq 10
$$

Here $\mathrm{L}_{\mathrm{t}-l}$ is the relative value which shows the difference between the target level (TL) and the present level, and $\mathrm{K}_{\mathrm{t}-l}$ is the relative value which shows the trend in the index of abundance. The parameters $g$ and $h$ are so called 'control gains'.

The procedure incorporates an algorithm to move TL closer to MSYL. Therefore, it is capable of automatic adjustment should there be changes of MSY\% and MSYL due to a change in the ecosystem, etc.
In order to prevent a drastic decrease of the stock size, the value of $g$ is made changeable according to the stock level. In addition, the range of variation of catches is constrained to be from $-20 \%$ to $+10 \%$ in order to preclude both large fluctuation of the catches and depletion of the stock.

Although the basic information required for implementation of the procedure is only the relative index of abundance $\left(X_{t}\right)$, it also utilises other information such as the age at recruitment ( $l$ ). It should be noted, however, that this procedure is not sensitive to estimation error in the age at recruitment. To determine $\mathrm{C}_{0}$, the constant catch taken over the first ten years, and $\mathrm{TL}_{0}$ at the start of the application of the S-T procedure, it is necessary to know the absolute value of the stock size at that time. As noted above, other information such as values of M, MSY\% and MSYL are not required.
When the management procedure is applied to an unexploited population and the index of abundance has no bias, robust management is achieved without the possibility of extinction. In almost all cases, the mean of the minimum population size detected in 100 simulations was more than $50 \%$ of the carrying capacity (K). When the index of abundance is proportional to the square root of the population size, the distributions of the population abundance in 100 simulations were almost the same as for the case where the index of abundance is proportional to the population size. Even in the case where the population has been depleted to 0.3 K and management commences without using any prior data, this management seemed to be successful. If the index of abundance was observed in every fifth year instead of every year, the population becomes stable at a higher level than the MSYL.

In short, the strength of the procedure is that it requires minimum information while utilisation of the stock is achieved without causing marked population decline and with relatively little fluctuation of the catches.

The disadvantage presently identified in this procedure is that it has to determine $g, h, C_{0}$ and $\mathrm{TL}_{0}$ for its implementation; its performance depends on these parameters. A number of simulation trials have identified that an appropriate value of $g$ is much greater than 1 , and $h$ is much smaller than 1 , i.e. this procedure has awarded heavier weight to changing the trend in stock size than to the stock level. The basic values of $\mathrm{g}=3$ or $5, \mathrm{~h}=0.04$ are set and $g$ is increased to twice or four times its size when the stock level declines.

An absolute abundance estimate $P_{0}$ is used to obtain $C_{0}$ and $\mathrm{TL}_{0} . \mathrm{C}_{0}$ is calculated as $\mathrm{C}_{0}=\mathrm{aP}_{0}$. The parameter a is set at $2 \%$, but for additional safety it might be better to adopt a value of $1 \%$. Information on K is required for setting the value of $\mathrm{TL}_{0}$. Once K is known, then $\mathrm{TL}_{0}$ is obtained by setting $\mathrm{TL}_{0}=\mathrm{bK}$. Setting the parameter b to 0.6 would be appropriate. Although some difficulty exists in estimating $K$, an approximate value of $K$ could be obtained by using $\mathrm{P}_{0}$ and the past catches.

An algorithm in which TL is moved close to MSYL,
regardless of $\mathrm{TL}_{0}$, has been incorporated into the procedure. However, it is desirable to set $\mathrm{TL}_{0}$ as precisely as possible, since the performance of the above algorithm mentioned is sometimes not efficient. Further study may be necessary regarding improvement of this algorithm.

Minimising the possibility of stock depletion on one hand and maintaining a high level of catches on the other are two ambivalent objectives of management. A procedure can only be formulated by giving heavy weight to one of these objectives. The S-T procedure has adopted utilisation without the risk of drastic decline of the stock as its primary objective.

Because of this primary objective, the application of this procedure tends to maintain high stock levels while keeping catch levels low. TL sometimes became much higher than MSYL. This can be improved to some extent by modifying the algorithm to adjust TL. It would be possible to achieve higher average catches, at the price of admitting the possibility of stock decrease, by changing the values of $\mathrm{g}, \mathrm{h}, \mathrm{C}_{0}$ and $\mathrm{TL}_{0}$, and the range of catch variation.

## Annex D3. THE DE LA MARE PROCEDURE

The procedure is designed to achieve the same basic operating objectives as the New Management Procedure (NMP), i.e. (1) stocks are exploited only if they are above a protection level (and thus the procedure aims to restore depleted populations to near MSY level as soon as possible) and (2) to stabilise the population at the level where the yield is expected to be at $90 \%$ of MSY.

There are two ways of setting catch limits of $90 \%$ of MSY. The first way is to try to estimate MSY directly by some means. This is the approach which has been tried since the NMP was first introduced. The second way is to find out what level of catch will maintain the population at the level where the yield would be expected to be about $90 \%$ of MSY. This management procedure is designed around the second approach. It does this by setting a target level of depletion (T) for the population at $75 \%$ of the unexploited initial level (strictly speaking the NMP would lead to a target of $74 \%$ using the conventional baleen whale model). The protection level (Q) is set at $55 \%$ of the unexploited initial population size (the NMP has $54 \%$ ). The slight differences in target population level and protection level have no significance, as the ones in the NMP are the result of an arbitrary choice of a model of production from a whale stock. The slightly different levels were chosen as a gentle reminder that they are deliberate choices, not based on certain knowledge of whale stock dynamics, but rather on a long held conjecture about what those dynamics might be. However, there is no reason why the levels chosen could not be the same as arising from the conventions adopted for the NMP.

A feedback approach is used to stabilise the stock at T, based on estimates of stock depletion. If the stock is estimated to be below T, the catch limit is reduced towards an amount below the estimated replacement yield (RY). If the stock is estimated to be above T, catch limits are increased towards an amount above RY. The relationship between the estimate of depletion ( $\hat{\mathrm{D}}$ ) and the proportion


Fig. 3. The control law for the de la Mare procedure.
of the estimate of replacement yield to be used as a catch limit is shown in Fig. 3. The formula for the graph is given by:

$$
G= \begin{cases}\frac{\hat{D}-Q}{T-Q} & \hat{D}>Q  \tag{1}\\ 0 & \hat{D} \leq Q\end{cases}
$$

The catch limit is given by:

$$
\begin{equation*}
\mathrm{C}=\widehat{\mathrm{RY}} \cdot \mathrm{G} \tag{2}
\end{equation*}
$$

However, the control law or decision rule illustrated in Fig. 3 is modified to take into account the inevitable uncertainty in estimates of depletion. If the estimated depletion has a lower confidence interval (at a chosen level of statistical significance) above the protection level Q then it is unlikely that the true stock is below the level where it should have been protected. If the lower confidence interval is below Q then the probability of inadvertently continuing exploitation increases. The latter situation can be dealt with by reducing catch limits to the extent required
to allow the population to increase to a level where the chosen lower confidence interval coincides with $Q$. This is achieved by reducing the slope of the control law through replacing $T$ in formula (1) with a value $\mathrm{T}^{\prime}$, determined from the following formula:

$$
T^{\prime}= \begin{cases}T & X \leq T-Q  \tag{3}\\ Q+X & X>T-Q\end{cases}
$$

where X is the difference between the estimated depletion and its lower confidence bound. In trials to date the $95 \%$ lower confidence bound has been used.

The catch limits are further modified by limiting the year to year fluctuation in catch limits by the following rule for the case $\mathrm{C}_{\mathrm{t}-1}>0$ :

$$
\begin{align*}
C_{t}=C_{t-1} * \alpha & \text { if } C_{t}>C_{t-1} * \alpha \\
C_{t-1} * \beta & \text { if } C_{t}<C_{t-1} * \beta \\
C_{t} & \text { otherwise } \tag{4}
\end{align*}
$$

Currently trials have used $\alpha=1.1, \beta=0.9$
However, if: $\quad \hat{D}<Q$ then $C_{t}=0$
If $\mathrm{C}_{\mathrm{t}-1}=0$ then the following rule applies:

$$
\begin{array}{ll}
C_{t}=G * \hat{R Y} & G * \hat{R Y}<\gamma \\
C_{t}=\gamma & G * \hat{R Y} \geq \gamma \tag{6}
\end{array}
$$

In the trials so far $\gamma=10$.
In the current implementation, $\hat{D}$ is estimated by fitting the baleen whale production model to the time series of abundance data. Estimates of MSY are not used.
In the trials presented to the Workshop, the estimation procedure used for depletion and RY varies with time during the trials to allow more flexibility in model fitting as more data become available. This helps to maintain the robust properties of depletion estimates over the whole simulation span, without leading to unnecessarily wide confidence intervals for depletion estimates during the early years when the data are few.

The most important feature of the overall results of the trials is that the management procedure has been quite successful in maintaining stocks at, or restoring them to, levels above the protection level. This is achieved even in the cases where estimates of abundance have high coefficients of variation, and are only collected at five yearly intervals. However, realising the full potential of the stocks depends to a much greater extent on the variance and frequency of the estimates of abundance. High coefficients of variation and longer inter-observation periods lead to reductions in catch limits.
The results of these trials show that reliable management is more difficult to achieve in rehabilitating a stock from a depleted state. With only five yearly abundance estimates having a CV of 0.4 , the results are very poor in terms of stock utilisation.

## Annex D4. THE COOKE PROCEDURE

The data requirements for this procedure are sightings estimates of abundance. The procedure does not use CPUE or biological parameters. A sightings estimate less than ten years old must be available to allow a catch.

The current population size is estimated independently of the overall status of the stock. Sightings estimates from the last ten years, if any, are averaged. To introduce an element of conservatism, the lower $95 \%$ confidence limit of this averaged estimate is taken. This is then further adjusted downward by the amount of catches taken in the years following the year to which the average refers.

A simple logistic population model is fitted through the above adjusted estimate of current population size: $\mathbf{N}_{\mathrm{t}-1}=$ $\mathrm{N}_{\mathrm{t}}-\mathrm{C}_{\mathrm{t}}+\mathrm{r} \mathrm{N}_{\mathrm{t}}\left(1-\mathrm{N}_{\mathrm{t}} / \mathrm{K}\right)$. Since the current population size has been estimated directly, then once $r$ has been determined, K can be calculated. The r parameter is estimated by fitting to the set of previous stock estimates, regarding them as relative indices of abundance only. In principle the maximum likelihood estimate of $r$ is selected. However, this would yield unstable estimates of $r$ in the early years when there are few data. Therefore, the data are combined with a prior likelihood function for $r$, which has the effect of tending to yield an $r$ value close to the present "anchor" value in the early years, but a value which is close to that suggested by the data in the later years. Having estimated $r$, the initial population size $K$ is calculated and the depletion of the stock (ratio of current
to initial stock size) is estimated. The raw catch limit is set at a certain proportion of the current stock estimate as determined by the control law. The control law is linear, ranging from zero at a stock current-to-initial ratio of $50 \%$ up to $\mathrm{r} / 2$ at the undepleted stock level. To reduce the annual fluctuations in catch limits, the actual catch limit is set as the average of this raw catch limit and the catches of the previous four years.

The primary aim of the procedure is to ensure that there is little risk of excessive depletion of any stock, and that a stock already depleted is allowed to recover to productive levels before exploitation recommences. The secondary aim is to ensure stable catches reasonably close to the MSY provided that the stock is not already too depleted. The simulation results show that the risk of excessively depleting an initially unexploited stock is negligible, and that depleted stocks are always allowed to recover to healthy levels before exploitation re-opens. For an initially unexploited stock, stable catches are achieved for the first 100 years, but the method clearly has difficulty handling data series longer than 100 years and can produce erratic variations in quota in the second 100 years. The potential yield is fully utilised when the true MSY is low, but an average annual catch of only about half the MSY is achieved when this is large. There is a tendency to wait longer than necessary before re-opening exploitation of an initially depleted stock.

## Annex D5. THE MAGNUSSON-STEFANSSON PROCEDURE

The essence of the method is a feedback law using the relative slope of a CPUE series (or some other relative abundance index) to stabilise a stock, together with jumps of a fixed percentage (e.g. 20\%) in the catch limit, whose purpose it is to move the population and catch towards the top of the sustainable yield curve (MSYL).

The procedure is as follows: after the feedback law has been applied for some time and a stability criterion is met, a jump in the catch quota is made (an increase or a decrease depending on the past history) and the feedback law switched on once again (see Fig. 4).


Fig. 4. Intersections of the net recruitment curve with three catch curves giving the equilibrium states $(\mathrm{Pi}, \mathrm{Ci})$, obtained by increasing or decreasing the catch by $100 \mathrm{p} \%$.

The stability criterion used in the simulations carried out so far is that the relative slope of the CPUE series should be less than $0.1 \%$ per annum.

If it is known initially that the stock is well above MSYL, then an increase in the catch limit is made every time the
stability criterion is met. This is continued as long as the average catch in the last ten years prior to a jump is increasing. As soon as this average is lower than the corresponding average prior to the last jump, a decrease is made in the catch limit and such decreases are continued until the average goes down once more.

On the other hand, if a population is depleted (or suspected of being close to MSYL or below), then the process is reversed, i.e. decreases in catch limit are made until the average catch ten years prior to a jump begins to decrease.

One potential advantage of this procedure is its simplicity; it is model independent in that it makes no assumptions about a population model, and thus no parameter estimation is required. No target level has to be set, but the probing strategy of the process gradually moves the population in the direction of the optimal level with little or no danger of extinction.
Furthermore, only the last 20 years or so of data in the CPUE series are used. Thus, the method is not sensitive to long term changes in catchability and/or changes in the age at recruitment which will distort the CPUE-population size relationship.

A possible disadvantage is that the progress towards MSYL is rather slow and the high variability in the relative abundance index might have the effect of making the jumps occur too rarely, thus further slowing down the progress towards the MSYL.

These disadvantages might be overcome to a certain extent by relaxing the jump criterion. One possibility might be to make a jump at least every $20-30$ years. Maintaining constant catches for a few years after a jump before switching on the feedback law would also speed up the progress towards MSYL.

# Annex $\mathbf{E}$ <br> Suggestions for Simple Progressive Management Strategies 

J. Shepherd

It is suggested that it would be desirable to test management procedures which:
(a) depend on the estimation of as few population parameters as possible; and
(b) are progressive, i.e. yield catch limits which change smoothly as a function of estimates (including population size), and avoid 'switching' from one rule to another at (more or less) arbitrary levels of uncertain parameters.
Two strategies are suggested - both could be implemented as simple modifications to Cooke's procedure. They are:
(1) constant fishing mortality (in the short term);
(2) a 'Cautious Management Procedure' (after Shepherd, 1981, Math. Biosci. 55: 179-87) in which fishing mortality is a linear function of stock size.
These are illustrated in Fig. 1. The Constant F strategy leads to catch limits which are linearly dependent on estimated stock size, whilst those for the Cautious strategy are quadratically dependent on stock size. These catch limits are similar to, but smoother than, those derived from other control laws (de la Mare, Cooke, Punt-Butterworth).

The constant $F$ strategy will lead to stabilisation at some non-zero population level, provided the level of $F$ chosen is not more than about double the true MSY\% level. This level may however be well below MSYL, and if the level of $F$ is too high, the stock will ultimately collapse; although if $F$ is a few percent, it would only do so very slowly declining by $1-2 \%$ each year. Nevertheless, to avoid this


Fig. 1. Constant F and cautious management strategies.
undesirable outcome, it would be preferable to attempt to select the $F$ level on the basis of some estimate of the population parameters (it should be set to about $\mathrm{r} / 2$ ). Such a procedure may already be sufficiently conservative in practice - this can only be established by testing. Note that one needs to estimate only current (absolute) population size, while (if $F$ is to be varied) the value of the product r.K is irrelevant.

For extra security, the Cautious procedure of Fig. 1 may be preferable. This will stabilise at a non-zero population level for any level of MSY\%, and this level should almost always be more than 0.5 K . It requires an estimate of K , in addition to that of $r$ (but both are available simultaneously from Cooke's method of estimation anyway).

It is suggested that these procedures could be implemented as a modification of Cooke's method set out in the following.
(1) Start with a prior distribution for $r$ (probably vaguer than that of Cooke in SC/F89/M6 - log normal for r, with median 0.04 , and log standard deviation of about 1.0 , so that values of $r$ about half or double the median are still quite likely). Use Cooke's method to estimate current population size (but preferably use the estimate itself rather than the lower $95 \%$ confidence limit). Start with a constant $F$ strategy (with $F=0.02$ ), and update $F$ to be $r / 2$ as estimates of $r$ (conditioned by the prior) become available.
(2) For the Cautious strategy, the procedure is the same, except that $F$ is set to be $(r / 2)(\mathrm{K} / 2)$ using the values of $r$ and K when these are available.
(Note: K should be a stabilised estimate, shrunk towards some prior and consistent with the final estimate of $r-i t$ is not quite clear whether Cooke's method does this already.)
(3) In both cases the quota is set to FP, without cut-offs for protection or anything.
I hope this is (a) comprehensible and (b) useful! Note that this approach is intermediate between the purely empirical methods, and those which are more model-dependent, but procedurally very close to that of Cooke. The underlying philosophy is that simplicity and smoothness are a 'Good Thing'. Whether such procedures are adequate (there is no reason to expect them to be optimal) is a matter for testing. They may at least provide a reference mark against which other more complex procedures (with more argument-generating free parameters) can be judged.

# Annex $F$ <br> On the Stability of Quota Control Systems 

J. Shepherd and J. Horwood

We have been worried for some time that quota control algorithms such as the NMP (and derivatives thereof) which switch fairly quickly between finite and zero catches as a function of (estimated) population size may not be very stable in practice. This could happen because such control algorithms have (in control theoretic terms) high gain, and using high gain in time delayed systems usually leads to instability. Magnusson and Stefansson (SC/F89/M9) find just such instabilities with their empirical control algorithm if too high a gain is used.

The problem can be studied crudely by examining the evaluation of (small) perturbations around the equilibrium state, as set out below. The population evolves according to:

$$
P_{t-1}=S P_{t}-C\left(P_{t}\right)+R\left(P_{t-k}\right)
$$

where $C\left(P_{t}\right)$ defines the catch limit algorithm and $R\left(P_{t-k}\right)$ the recruitment function. At equilibrium, $\mathrm{P}_{\mathrm{t}+1}=\mathrm{P}_{\mathrm{t}}=\mathrm{P}^{*}$ and thus:

$$
(1-\mathrm{S}) \mathrm{P}^{*}+\mathrm{C}\left(\mathrm{P}^{*}\right)-\mathrm{R}\left(\mathrm{P}^{*}\right)=0
$$

Expanding about this equilibrium point, and writing $\mathrm{x}_{\mathrm{t}}=$ $P_{t}-P^{*}$, leads to the dynamic equation for small perturbations about equilibrium:

$$
x_{t+1}=S x_{t}-\frac{d C}{d P} x_{t}+\frac{d R}{d P} x_{t-k}
$$

where the differentials are evaluated at the equilibrium point. We seek (potentially oscillatory) solutions where $\mathrm{x}_{\mathrm{t}+1}=\mathrm{Z}_{\mathrm{xt}}$. These exist for:

$$
\mathrm{Z}=\left(\mathrm{S}-\frac{\mathrm{dC}}{\mathrm{dP}}\right)+\left(\frac{\mathrm{dR}}{\mathrm{dP}}\right) \mathrm{Z}^{-\mathrm{k}}
$$

The full analysis of equations of this type is tedious (see Horwood, J.W. and Shepherd, J.E., 1981. The sensitivity of age-structured populations to environmental variability. Math. Biosci. 57:59-82). However, even for $\mathrm{k}=0$ (no time lag in recruitment) unstable solutions occur if:

$$
\frac{\mathrm{dC}}{\mathrm{dP}}-\frac{\mathrm{dR}}{\mathrm{dP}}>1+\mathrm{S}
$$

and the instability of such systems usually gets worse as the time lags increase. Thus trying to apply a sharp control on catches (i.e. dc/dp large) while on the right hand side of the recruitment curve $\mathrm{dr} / \mathrm{dp}$ becomes relatively large and negative, is a 'Bad Thing' from the point of view of establishing stable control.

This analysis applies only to linear control of small perturbations. In practice most of the controllers being tested may be quite non-linear, and may be asked to deal with large perturbations. This almost certainly makes things worse - maybe much worse. It is for this reason that progressive control algorithms are probably a 'Good Thing'.
Note that what one really needs to study is the response of the controlled system to noise, and that for the whaling problem this is mainly due to errors in the control (catch limit) itself, because of mis-estimation (e.g. of population size) due to observation errors.

It is suggested that this matter warrants further study, and that meanwhile some simulation testing of more progressive control algorithms would be desirable.

## Annex G Possible Questions to Address in Second Stage Screening

## Prefix notation

* Important and readily implemented
** Important and should be implemented in later second stage screening if possible


## Mistakes in models

* M1 Form of model (see Tests 1a and b, Section 8).
* M2 Time variable parameters (see Tests 2 a and b, Section 8).
** M3 Spatial effects in density dependence.
** M4 Variable recruitment/survival (i.e. underlying dynamics are not completely deterministic).


## Problems with data

* D1 Bias in absolute abundance data (see Tests 3a and b, Section 8).
** D2 Changes in assessment methods, e.g. changed survey methods with different bias.
D3 Autocorrelation in time series.
* D4 Trends in whaling practice, e.g. changing age at recruitment, introduction of new technology (see Tests 4a and b, Section 8).
** D5 Density is not uniform and whaling activity is not random, and/or whaling moves from one place to another.
D6 Key data cease to be collected.


## Problems with stock identity

* S1 Over or under subdivision of stocks (see Tests 5a, $5 b$ and 6, Section 8).
** S2 Stock distributions are geographically variable.
** S3 Concentration of effort at stock boundaries.
** S4 Whaling is geographically variable on underdivided stock.


## Ecological interactions

** E1 Competition, predator-prey relationships, mutualism.
E2 Trends in environment
Distribution of whaling activity in stock unit
** A1 Same stock is exploited in more than one place, e.g. in breeding grounds, on feeding grounds, or while migrating.
** A2 Whales do not mix rapidly throughout stock range, site fidelity.
** A3 Coastal and pelagic whaling.

## Innovation

I1 In whaling technology.
(*) I2 In assessment methodology (see D2).
I3 Calibration.

## Costs of management

** C1 Trade-off between cost of management and size of catch.

## Annex H Error Structure for CPUE Data

Two processes contribute to fluctuations in CPUE as an index of population size. For simplicity of explanation, consider the situation where CPUE is proportional to population size, with catchability (q) defined as the constant of proportionality. Then these two processes are as set out below:
(1) Catchability (q) fluctuations from one year to the next - predominantly the result of changes in environmental conditions. The typical size of the fluctuations is independent of the size of the catch.
(2) Sampling error - catching is a sampling process and so gives rise to this error. The typical size of the fluctuations depends on the size of the catch, smaller catches leading to larger fluctuations.
The trials set out in the 1987 Workshop assumed that one or other of these two processes was completely dominant. Trials with a log-normal error distribution of fixed CV corresponded to catchability fluctuation dominance; trials with a gamma error distribution corresponded to sampling error dominance in a Poisson-like process, with the formula given in that report providing the distribution of effort for fixed catch and population size.

To reduce the number of trials required, a composite
error distribution has now been specified. For the case of CPUE proportional to population size, this is:

$$
\begin{array}{rlrl}
\mathrm{CPUE} & =\mathrm{qP} \mathrm{e}^{\varepsilon} & & \varepsilon \text { from } \mathrm{N}\left[0, \sigma^{2}\right] \\
\sigma & =\ln \left[1+(\mathrm{CV})^{2}\right] & & \text { i.e. }(\mathrm{CV})^{2}=\mathrm{e}^{\sigma^{2}}-1 \\
\mathrm{CV} & =0.2+0.2 \sqrt{\mathrm{MSY} / \mathrm{C}}
\end{array}
$$

Thus the error distribution is log-normal, but with a CV that is a function of the size of the catch taken. The first term in the equation for the CV corresponds to the catchability fluctuation component and is independent of the size of the catch made. The second term is a simple approximation to the gamma distribution representation of sampling error (see Kirkwood, 1981, Rep. int. Whal. Commn 31:729-35, who explains that for an overdispersed Poisson process, the distribution of C for fixed effort is approximately normal with a fixed variance).

Note that the formula for CV is such that the CV will be larger for smaller catches. When a catch equal to the MSY is taken, $\mathrm{CV}=0.4$ corresponding to the specification for the gamma distribution at the 1987 Workshop.

When CPUE is proportional to the square root of population size, the formulae above are unchanged except that:

$$
\text { CPUE }=\tilde{q} \sqrt{P} e^{\varepsilon}
$$

## Reviews and Studies <br> funded as part of the Comprehensive Assessment

# Survey Techniques for Estimating Abundance of Cetaceans 

A.R. Hiby and P.S. Hammond<br>Sea Mammal Research Unit, Natural Environment Research Council, clo British Antarctic Survey, High Cross, Madingley Road, Cambridge CB3 OET, UK

Accepted March 1989

ABSTRACT
Rather than try to provide an abstract for this extensive report, we feel it more appropriate to list the section headings.
(1) INTRODUCTION
1.1 Objectives
1.2 Some definitions
(2) SURVEY TECHNIQUES
2.1 Effective search range
2.2 Strip transect
2.3 Line transect
2.4 Methods based on measurement of observer behaviour
2.5 Methods based on survey of cues - 'Cue Counting'
2.6 Shipboard acoustic survey
2.7 Land-based visual survey
2.8 Land-based acoustic survey
(3) DATA COLLECTION
3.1 Estimating the position of each sighting relative to the platform
3.2 Sighting conditions
3.3 Independent observers
3.4 Survey platforms
3.5 Land-based visual surveys
3.6 Acoustic surveys
(4) DATA ANALYSIS
4.1 Estimation of $f(0)$
4.2 Estimation of sighting efficiency from independent observer experiments
4.3 Whales or pods as targets?
(5) MODELLING PERPENDICULAR DISTANCE DATA FROM LINE TRANSECT
5.1 Definition of the problem
5.2 Investigation of a 'Good' data set
5.3 Investigation of a 'Poor' data set
(6) EFFECTS OF WHALE MOVEMENT
6.1 Transects located independently
6.2 Effects of whale drift across a grid survey
6.3 Movement in reaction to platforms
(7) SURVEY DESIGN
7.1 Introduction
7.2 Simple random survey design
7.3 Variable coverage probability designs
7.4 Variance estimation
7.5 Factors to consider in the design of surveys
7.6 Allocation of survey effort
7.7 Non-random survey designs
(8) CUE COUNTING AND ESTIMATION OF SURFACING RATE
8.1 Estimation of abundance from CC data
8.2 Estimation of surfacing rates
8.3 Analysis of CC data from independent observer experiments

## 1. INTRODUCTION

This work was commissioned by the International Whaling Commission (IWC) as part of the Comprehensive Assessment Work Plan for 1986/87 (IWC, 1987). Our report has been guided by the terms of reference specified and thus our general objectives can most easily be stated by listing below the terms of reference.

[^5](1) Review the suitability (including the use that has already been made) of the following techniques for estimating current abundance of whale stocks and monitoring trends in abundance: land-based sightings; aerial sightings; shipboard sightings; land-based acoustic surveys; and shipboard acoustic surveys.
(2) Identify the critical assumptions of each technique, means of determining the extent to which these assumptions are satisfied, and any alternative procedures to be followed if critical assumptions are not satisfied.
(3) Advise on appropriate methods of analysis of data from each technique for estimating current abundance and determine the precision of those estimates.
(4) Consider the practical limitations of each technique for censusing particular species and stocks of whales.
(5) Provide a guide for non-specialist researchers to aid in the choice of an appropriate census technique for particular whale stocks and species to estimate with a given degree of precision: (a) current abundance; (b) trends in abundance.

These terms of reference encompass a very wide subject area. Our coverage of the various aspects of cetacean surveys varies considerably. Some aspects are dealt with mainly by reference to other work in the form of a brief review; others by development of a new idea. Our coverage also varies in its direct relevance to the Comprehensive Assessment and the terms of reference of the commissioned 'review'. We have concentrated on techniques that can be used for the survey of given stocks and species in the immediate future rather than with methods of analysis for existing data. This is partly because most existing data sets lack one or more elements which would make further analysis worthwhile, but also because we were aware of the acute need for the type of 'guide' mentioned in the terms of reference above. The Workshop on the Design of Sightings Surveys held in Seattle in 1980 (IWC, 1982), useful as it was for exchanging ideas and motivating further research, did not provide a comprehensive 'how-to-do-it' guide for researchers about to embark on a survey. An agreed methodology had not emerged at that time; it may have been that the procedures during shipboard sightings surveys were too similar, superficially, to those on scouting boats for the need for much theoretical input to be apparent. An agreed methodology has still not emerged, and so it is still not possible to produce a comprehensive guide. However, there have been significant advances in a number of areas, for example, survey design, modelling the distribution of perpendicular distance to sightings, and the measurement of angles and distances to sightings and our report does make specific recommendations in these areas. Where no specific recommendations on data analysis are possible we suggest that data be collected to allow alternative analyses to be performed.

For the purpose of this report we take 'survey' to mean the estimation through sampling of absolute or relative abundance, excluding methods based on catch statistics or marking (whether by natural or artificial marks). This then leaves all techniques based on 'sightings' and acoustics. We follow the approach introduced by Cooke (see Section 7) of estimating abundance directly (rather than via the concept of 'density') by taking a sample from the population and calculating the probabilities of inclusion in the sample. The survey techniques considered here can then be characterised as those where the probabilities of inclusion follow from the design of the survey - in mark/recapture methods the inclusion probabilities are calculated from the pattern of recaptures of recognisable individuals, and in removal methods from the changes in sample size following known removals. Although mark/recapture methods are excluded from consideration we do not wish to suggest that those methods and the ones that are considered are mutually exclusive. For example, it is possible to adopt a mark/recapture approach to analysis of data from independent observers in sighting surveys
(Buckland, 1986). It may also be possible to combine mark/recapture and 'sightings' data in certain cases (Hiby, 1987).

The term 'abundance' requires definition for each survey. To what 'targets' does the estimate apply, over what region and at what time? An example of abundance may be the number of minke whales along a given stretch of coastline, out to a given depth contour, in a given month. Alternatively, it may be the number of gray whales passing within a given distance of a point on the coast, during a given interval of time. Such estimates can then be accompanied by one or more inferences which, in general, become less useful as the spatial and temporal constraints become more severe. In the gray whale example, the estimate may apply to the entire population if the given distance from the coast is sufficiently large, and the given interval sufficiently long. In examples where the whales are surveyed from a moving platform, the quality of inference depends on the degree to which the given region encompasses the range of the species or stock at the time of the survey. If significant movement into or out of the region occurs during the period of the survey, the abundance estimate refers to a summary statistic of the number of targets within the region during that period, such as the mean number.

It is also important to consider whether an absolute or relative estimate (index) of abundance is of primary interest because, in certain cases, it may be possible to obtain an accurate estimate of relative abundance suitable for monitoring trends but only a rough estimate of absolute abundance (e.g. Best, 1985).

Although we have limited our report to the estimation of abundance from sampling surveys, it should be noted that such surveys can provide other types of information, for example on the distribution of targets in relation to oceanographic features (e.g. Payne, Nicholas, O'Brien and Powers, 1986; Dohl, Bonnell and Ford, 1986; Smith, Dustan, Au, Baker and Dunlap, 1986; Au and Perryman, 1985; Ljungblad, Moore and Clarke, 1986; and Moore, Clarke and Ljungblad, 1986.

## 2. SURVEY TECHNIQUES

This Section provides an introduction to the various techniques available, giving one or two examples of application of the technique to surveys for cetaceans, the types of data required, the derivation of the corresponding estimators and the assumptions made. A number of alternative ways of classifying these techniques could be constructed; however, the choice is not so large as to make this necessary or worthwhile, and they are therefore simply listed. For the same reason we have listed the strengths and weaknesses of each technique under the separate headings, instead of trying to produce a decision tree for choice of method.

Traditionally, survey techniques to estimate cetacean abundance from moving platforms have been based on the concept of estimating density via sampling of areas of sea. The problem is to obtain a representative sample from a highly inhomogeneous population (the population consisting, in this case, of units of sea area). Stratified sampling can help, but as pointed out by Burnham, Anderson and Laake (1980, pp.31-2) the potential for stratification is severely constrained by the need to obtain adequate sample sizes within each stratum. Uniform coverage probability is required within each stratum and
this may be difficult to obtain given the logistic constraints faced by the survey. In techniques based on counting behavioural cues, the problem is exacerbated by the need to ensure uniform coverage in terms of time as well as space, so that in practice the speed of the sighting platform cannot be allowed to vary while on effort. Another problem is that to estimate abundance, the estimate of mean density is multiplied by the area of the region to which the survey applies - that area may be difficult to calculate if, for example, the region is bounded by a rapidly changing ice edge as has occurred in the cruises for minke whales in the Southern Hemisphere undertaken as part of the IWC's International Decade for Cetacean Research (IDCR - see Section 2.3).

In our presentation of survey techniques, we follow Cooke (1986a; 1987a; Section 7) and place the emphasis on sampling the whales themselves rather than the sea. The population under consideration is therefore the total number of whales in the area of interest rather than a number of units of sea area. The estimate of abundance is obtained directly as a function of the probabilities of detection for the $n$ whales appearing in the sample; in the case of line transect sampling it is simply the sum of the reciprocals of these probabilities:

$$
\hat{\mathbf{N}}=\sum_{i=1}^{n} 1 / p_{i}
$$

Here, $p_{i}$ is the probability of detection for the $i^{\text {th }}$ whale in the sample and depends on the degree of 'coverage' in that area afforded by the chosen cruise track design, and on the ability of the sighting platform to detect whales. This paper is thus concerned with estimation of these factors.

To see that the estimator is unbiased for population size N , associate, with each whale in the population, an indicator variable $x_{j}$, which has value 1 if whale $j$ is included in the sample and 0 if it is not included. Then

$$
\hat{\mathrm{N}}=\sum_{\mathrm{i}=1}^{\mathrm{n}} 1 / \mathrm{p}_{\mathrm{i}}=\sum_{\mathrm{j}=1}^{\mathrm{N}} \mathrm{x}_{\mathrm{j}} / \mathrm{p}_{\mathrm{j}}
$$

and

$$
\mathrm{E}(\hat{\mathrm{~N}})=\sum_{\mathrm{j}=1}^{\mathrm{N}} \mathrm{E}\left(\mathrm{x}_{\mathrm{j}} / \mathrm{p}_{\mathrm{j}}\right)=\mathrm{N}
$$

because

$$
\mathrm{E}\left(\mathrm{x}_{\mathrm{j}}\right)=\mathrm{p}_{\mathrm{j}}
$$

This inverse-detection-probability approach eliminates the need to ensure uniform coverage probability within a stratum and does not require a measure of the area of the region to which the survey applies. It also facilitates the study of the effects of target movement (see Section 6) and the incorporation of information on behavioural cues into the abundance estimator (Section 8).

### 2.1 Effective search range

It is convenient to consider the problem of estimating detection probabilities in two parts: first, given the geographical location of a target, what is the probability that a transect will be placed within a given distance of that location and, second, what is the probability the target will then be detected? The first part of the problem concerns the design of the survey, the second the effective search range of the platform.

Intuitively the effective search range represents the ability of the sighting platform to detect targets. In the case of strip or line transect it is the effective search half-width (esw). Let $\mathrm{G}_{\mathrm{i}}(\mathrm{y})$ represent the probability density function for the closest approach of a transect to the location (assumed stationary) of the $i^{\text {th }}$ whale in the sample, and
$g(y)$ the conditional probability that a stationary target at perpendicular distance $y$ from the transect will be detected. The unconditional detection probability is then

$$
p_{i}=\int_{0}^{\infty} G_{i}(y) g(y) d y
$$

If the density $\mathrm{G}_{\mathrm{i}}(\mathrm{y})$ is approximately constant at $\mathrm{p}^{\prime}{ }_{\mathrm{i}}$ over the interval $(0, \mathrm{w})$ and any sightings beyond $w$ are excluded, the detection probability becomes

$$
p_{i}=p_{i}^{\prime} \int_{0}^{w} g(y) d y
$$

The integral $\int_{0}^{w} g(y) d y$ is the effective search half-width.
The equivalent quantity for the situation where behavioural cues are counted is defined in Section 2.5.

Much of the theory of sighting surveys concerns the estimation of effective search range, so it is important to be comfortable with the concept. One way to see why the integral of $\mathrm{g}(\mathrm{y})$ is the effective search half-width is to note that the number of whales missed within this distance is equal, in expectation, to the number of whales detected beyond it (Gates, 1969). Consider whales within distance w of the trackline, where $g(y)=0$ for $y>w$. The probability a whale is missed given it is within esw is:

$$
1-\int_{0}^{\mathrm{esw}}[1 / \mathrm{esw}] \mathrm{g}(\mathrm{y}) \mathrm{dy}
$$

and the probability it is seen given it is beyond esw is:

$$
\int_{e s w}^{w}[1 /(w-e s w)] g(y) d y
$$

But the number of whales within and beyond esw are in the ratio esw:w-esw, so the number missed within and seen beyond esw are in the ratio
i.e. 1:1.

$$
\mathrm{esw}-\int_{0}^{\text {esw }} g(y) d y: \int_{\text {esw }}^{w} g(y) d y
$$

Another way of motivating the definition of esw, which is more relevant to the presentation in this paper, is to note that $p_{i}{ }^{\prime}$ is equal to the probability that the transect falls within one distance unit of whale i or, in other words, the probability a strip of width two distance units covers whale i. Thus $p_{i}$, which is the detection probability for whale $i$, is the probability a strip of width 2esw distance units covers whale i. 2esw is therefore the total effective strip width.

### 2.2 Strip transect

Examples of strip transect surveys for cetaceans are Smith, Hamill, Burrage and Sleno (1985) and Leatherwood (1982).

In a strip transect survey the sighting platform (ship or aircraft) moves along each transect and records the number of targets ( n ) sighted within a predetermined distance ( $\mathrm{w)}$ of the trackline. In the case of aerial survey that distance can be identified by marks placed on wing struts or windows. If all targets within the strip are certain to be detected then, by the definition given in Section 2.1 above, effective search width (esw) is equal to the strip width used on the survey, $w$ (from the trackline to the strip boundary). The estimate of abundance is thus

$$
\hat{\mathbf{N}}=\sum_{i=1}^{n} 1 / p_{i}^{\prime} w
$$

Suppose the total track length is $L$ units, the area surveyed A square units and the track is designed to give uniform coverage. Then $\mathrm{p}_{\mathrm{i}}{ }^{\prime}$, the coverage probability based on a nominal strip width of two units, is 2L/A and

$$
\hat{\mathrm{N}}=\mathrm{nA} /(2 \mathrm{LW})
$$

which is the usual estimate of abundance derived using the concept of mean density.

The assumption that all targets within the strip are detected may be weakened to the assumption that the probability of detection is constant across the strip at, say, $\mathrm{g}(0)$. In this case the esw is reduced to $\mathrm{g}(0) \mathrm{w}$. It may be possible to estimate $g(0)$ by using two sets of observers. Each set records its own sightings and is kept unaware of the sightings made by the other set. 'Duplicate' sightings made by both sets are identified by a third person in real time or at the stage of data analysis, and the duplicate proportion used to estimate $g(0)$ (see Sections 3.3 and 4.2). However, the estimator is based on the assumption that the probability of a target being sighted is independent of whether it is sighted by the other set. The true proportion of whales missed cannot be estimated if some whales are unavailable to both sets of observers because they were on a long dive when the platform passed by. Estimation of that proportion would require some sort of ground truth exercise or a modification of the sampling strategy to permit calibration using behavioural measurements (see Section 2.5).

The sighting targets may be individual whales, pods or schools. A pod or school is included in the sample only if its 'centre' lies within the boundaries of the strip. Because strip transect sampling is most likely to be carried out as part of an aerial survey, temporary diversions from the trackline to check species identification or to estimate school size (possibly by using photography e.g. Scott, Perryman and Clark, 1985) are unlikely to lead to the biases incurred as a result of 'closing mode' procedures in shipboard surveys (see Section 2.3).

Target motion, more fully considered in Section 6, can affect shipboard strip transects in a number of ways. The effects will be negligible for aerial survey. From ships, random target motion can cause an increase in sighting rate as a result of targets entering the 'sides' of the strip. In effect the value of $w$ used for esw is an underestimate in this case. One way to overcome that problem is to include only those targets within the strip when abeam of the platform. That modification will not, of course, eliminate any bias resulting from targets being attracted or repelled across the strip boundary as a consequence of reaction to the ship (see Section 6.3).

It is difficult to conceive of a situation where strip transect sampling would be recommended in place of line transect. Burnham and Anderson (1984) provide convincing arguments for collecting distance data in any survey. In some circumstances there may be no alternative, such as if sightings are to be obtained from platforms of opportunity and the measurement of perpendicular distance to each sighting is not feasible.

### 2.3 Line transect

Line transect sampling can be considered as a generalisation of strip transect and a number of the comments of the last section apply to this section also. Examples of the application of line transect sampling to aerial survey for cetaceans are Hay (1982), Scott and Gilbert (1982), Smith (1981) and Holt and Powers (1982).

The series of IDCR cruises designed to assess the abundance of minke whales in Antarctic areas serves to illustrate the application to shipboard survey. The development of the survey techniques and experiments designed to test assumptions are reported in a number of papers, for example Butterworth, Best and Basson (1982),

Butterworth, Best and Hembree (1984), the cruise reports and the reports of the IWC Scientific Committee's sub-committee on Southern Hemisphere Minke Whales (e.g. IWC, 1986b). The monograph by Burnham et al. (1980) provides a comprehensive review and treatment of the general theory of line transect sampling and is a useful starting point for the consideration of line transect sampling in cetacean surveys. Hammond (1986a) reviews the application of line transect sampling to dolphin populations. Radial distance models for line transect data are considered by Hayes and Buckland (1983). A methodology for calculating the effect of target motion on sighting rate was developed by Yapp (1956) and Skellam (1958). The work by Koopman (1980) on the theory of search provides a number of results relevant to surveys for cetaceans, for example, the treatment of detection probability in terms of the hazard rate. The models for hazard rate in terms of target size, and intrinsic and apparent contrast, are also relevant to the consideration of 'weather' factors experienced during a survey, although this aspect continues to be largely neglected.

The basic data requirements are the number of targets sighted from the platform over a measured distance on sighting effort, and an estimate of the perpendicular distance $y$ from the trackline to the position of each target at the moment it is sighted. Normally the targets are pods or schools; species are confirmed and pod or school size is estimated from the trackline ('passing mode') or by temporary diversion from the trackline to the sighting ('closing mode'). In aerial survey, the aircraft delays closure until the sighting comes abeam, when y can be measured, and rejoins the trackline at the point from where it left it. In shipboard survey other strategies may be deployed; for example the ship may close with the sighting immediately and rejoin the trackline by following a convergent course. Sightings made while the ship is closing with and confirming the species composition and pod size of the sighting are called 'secondary' and those made from the trackline 'primary'. Only a subset of sightings may be confirmed, for example those lying within a certain distance of the trackline. When pods or schools are the targets, the estimated distance $y$ is to the perceived 'centre' of the pod or school. It is possible for some ambiguity about pod definition to occur, for example, which of the whales seen on closure to include in the primary sighting, and which to designate as secondary (e.g. Hay, 1982).

A number of possible biases are associated with closing mode surveys. Horwood (1981) pointed out that effort in high density areas could be reduced because of the larger proportion of the transect not covered on effort. Kishino and Kasamatsu (1987) have investigated this effect using simulation modelling. This suggests that the platform should return to the point on the trackline from where the confirmed primary sighting was made. This could, however, lead to an upward bias in areas of high density due to a stop/start effect. As a result of random whale movement the area ahead of the vessel, effectively 'cleared' of targets while on effort, is filled up during the time used for confirming the sighting. This bias is a function of whale density; simulation study suggests an upward bias of about $35 \%$ at a density of 1 pod per n.mile ${ }^{2}$. It can be eliminated by ensuring adequate 'secondary' search effort prior to the recommencement of primary search effort to 'clear' the area of potential sightings.

The following derivation of the esw is due to Seber (1982). Let $w$ represent a convenient upper truncation
value for y ; all or most sightings have y values less than w . Assume that a stationary target within distance $w$ of the trackline is equally likely to be at any distance $y$ from the trackline, that is, the probability density for y is uniform at $1 / w$. This results from random transect placement if $w$ is 'small' on the scale of the region to which the survey applies. Then the probability the whale is detected is

$$
(1 / w) \int_{0}^{w} g(y) d y
$$

The probability the whale is seen at $\mathrm{y}<\mathrm{y}^{*}$ is

$$
\left.(1 / w) \int_{0}^{y^{*}} g(y) d y\right]
$$

and the probability the whale is at $y<y^{*}$ given it is seen is

$$
\begin{gathered}
{\left[(1 / w) \int_{0}^{y^{*}} g(y) d y\right] /\left[(1 / w) \int_{0}^{w} g(y) d y\right]} \\
\int_{0}^{y^{*}} g(y) d y / e s w
\end{gathered}
$$

Therefore the probability density function (pdf) for the perpendicular distance to targets sighted within $w$, is

$$
f(y)=g(y) / \text { esw }
$$

Thus esw $=g(y) / f(y)$ and in particular esw $=g(0) / f(0)$.
Thus, given $g(0)$, the esw can be estimated by fitting a function $f(y)$ to the observed distribution of perpendicular sighting distances and evaluating the fitted curve at $\mathrm{y}=0$. Section 5 considers the choice of appropriate function for $\mathrm{f}(\mathrm{y})$.

In the use of the line transect method for estimating abundance of terrestrial plants or animals, the assumption $g(0)=1$ may be justified in many cases. In application to cetacean surveys that assumption is rarely justified except perhaps for those species forming large schools with a number of animals visible at the surface at all times. Even then, the assumption has to be made that all schools are larger than a minimum threshold size (e.g. Holt and Powers, 1982). For other species, shipboard surveys may approach the condition $\mathrm{g}(0)=1$, depending on pod size, surfacing rate and sighting conditions but for aerial surveys, $g(0)$ will certainly be less than 1 . The use of the $\mathrm{g}(0)=1$ assumption will therefore lead to upward bias in estimation of esw and downward bias in the estimation of abundance. The size of that bias will depend on the sighting conditions and the observers on the survey, and as a result the value of the abundance estimator as an index to monitor change over time is questionable. Some researchers have investigated the effect of various measures of sighting condition on sighting rates (e.g. Scott and Gilbert, 1982; Leatherwood and Reeves, 1982; Gunnlaugsson and Sigurjónsson, 1989) but the results do not permit the degree of calibration required to produce a reliable index.

The general problem of the application of line transect methods to cetaceans is, then, estimation of the ratio $g(0) / f(0)$, requiring extra data in addition to the number of sightings and perpendicular distances $y$. The IDCR Southern Hemisphere minke whale assessment cruises document the progress made in addressing this problem. The use of varying platform speed to provide an estimate of $g(0)$ was investigated by Butterworth, Best and Basson (1982), the concept of the 'hazard rate' being used to derive the estimator. That approach was eventually rejected in favour of the use of 'independent' observers, as described
for the strip transect case. Platforms have now been constructed which are intended to be visually and acoustically isolated from other observers on the vessel. The current procedure is to carry out line transect sampling from these positions throughout the passing mode sector of the cruise (e.g. Butterworth and Borchers, 1988). This is in contrast to the earlier procedure, carried out during experimental periods only, where the two sets of observers were on different ships which moved along parallel tracks separated by 0.3 or 1.0 n.miles (e.g. Butterworth et al., 1982; Butterworth, Best and Hembree, 1984; Buckland, 1987). (In analysis of IDCR data the product 'eh' has been used. This is synonymous with $f(0) \bar{y} / g(0)$ the factor $h$ representing $1 / g(0)$ and e representing $f(0) \bar{y}-e$ was used as a correction factor accounting for the difference between the actual value of $f(0)$ and $1 / \bar{y}$, which is the maximum likelihood (ML) estimate for $f(0)$ if $f(y)$ is the untruncated negative exponential.)

The identification of duplicate sightings remains problematic, but the independent observer method, more fully described in Section 3.3 and Section 4.2, may have the potential for robust estimation of $g(0) / f(0)$ in the case of shipboard line transect sampling. The potential is limited in the case of aerial line transect surveys because, as explained for strip transect, the requirement for sighting probabilities to be independent is violated when some whales are invisible (below the surface) as the platform passes.

The observations concerning target motion from Section 2.2 apply to line transect also, with the added complication that for shipboard surveys, the estimate of esw may also be affected due to distortion of the frequency distribution for perpendicular distances to sightings (see Section 6). This will certainly affect estimates of $1 / f(0)$, although estimates of $g(0) / f(0)$ from the independent observer procedure may well be robust to this type of distortion. One way of avoiding any possible bias is to record distances at which sightings pass abeam of the ship (e.g. Whitehead, 1982).

The estimate of esw is a statistic of the data and therefore has an associated variance. In the case where detection probabilities for each target encountered are independent (note: this is not required for unbiased estimation) and depending on the chosen functional form for $f(y)$, an estimate of the variance may exist in closed form (Burnham et al., 1980, pp.51-5) or be based on the asymptotic properties of ML estimators. If detection probabilities are not independent, for example if whales rather than pods are treated as the targets, then variance has to be estimated using jackknife or bootstrap techniques.

### 2.4 Methods based on measurement of observer behaviour

 In the method of Doi (Doi, Kasamatsu and Nakano, 1982) the line transect approach was modified to include estimation of $g(0)$ based on measurement of observer behaviour. The probability of detecting a blow from a minke whale was modelled as a function of its distance from the vessel, the angle of view of the binoculars used by topmen (topmen on the IDCR surveys search almost entirely using binoculars), the rate of sweep of the binoculars and the duration of the blow. The probability of detection for a blow occurring within the field of view of the binoculars was assumed to depend on the rate of sweep of the 'vision line' across the blow, rw, where $r$ is radial distance to the blow and $w$ is angular velocity of the binocular sweep. The resulting function was fitted to thedistribution of radial distances to sightings ahead of the vessel. An estimate of average inter-blow interval for Antarctic minke whales was then incorporated to give a measure of $g(0)$ (or kw in the notation of the Doi model). Measurements of binocular sweep rate were obtained using a video camera sited above the barrel to look vertically down onto the topmen.

The method is in the spirit of the statement by Koopman (IWC, 1982, pp.543-4) which suggests, in essence, that reliable estimation of esw cannot be obtained by curve fitting alone and should be based on the physical structure of the sighting process. However, Doi's approach falls short of estimating esw and as a result misses the opportunity of comparing the two dimensional distribution of sighting positions predicted by the model with that observed during the survey, as implicitly recommended in Koopman's statement.

A second approach based on measurement of observer behaviour was suggested in Hiby and Thompson (1985) and Hiby (1985). The method incorporated a specified degree of random whale movement and estimated an esw defined accordingly; that is $g(y)$ was defined as the probability of detecting a whale destined to pass abeam at distance $y$, not the probability of detecting a whale stationary at distance $y$ from the trackline. The esw was estimated by determining the hazard rate using sighting positions and the angular allocation of effort by the observers, measured using a mast-head camera, as described above. The hazard rate is a function $\mathrm{H}(\mathrm{r}, \theta)$ which gives as $\mathbf{H}(\mathrm{r}, \theta) \mathrm{dt}$, the probability that a target at distance r from the vessel and angle $\theta$ from the trackline is detected during the interval dt. It was assumed to be separable into $H(r) \times H(\theta)$ and $H(\theta)$ was assumed to be proportional to the angular allocation of effort. $\mathrm{H}(\mathrm{r})$ was determined by fitting to the distribution of observed sighting positions in $r$ and $\theta$.

The confidence limits on the estimate of esw by fitting to sighting positions only are too wide to be useful. However, data from the independent observer procedure can be incorporated into the technique which thus provides an alternative method of analysis. The method has a number of potential advantages, but requires further assessment. In terms of data collection, the only difference between this approach and line transect sampling is that estimates of $r$ and $\theta$, not just $y$, are required, and that measurements of the allocation of effort with angle should be obtained, if possible. Even in the usual analysis of line transect data, those measurements can be useful for checking assumptions.

### 2.5 Methods based on survey of cues - 'cue counting'

As mentioned in Section 2.3 the condition $g(0)=1$ for line transects will hold in cetacean surveys only under exceptional circumstances. For aerial survey, $g(0)$ may be much less than 1 and the independent observer procedure is of limited potential. Nevertheless, aerial survey has very significant advantages, in particular in its ability to integrate more effectively over the inhomogeneous spatial distribution of whales (reducing the dependence between $p^{\prime}$ i values for different targets). Furthermore, the effects of target motion and the logistic problems of surveying areas adjacent to indented coastlines or shifting ice-edges are eliminated. There is therefore a strong motivation to derive unbiased estimators of abundance from aerial survey data.

One such method is based on the idea that whales are missed along the trackline because they are on a dive as the platform passes; a whale on the trackline and continuously at the surface would be certain of detection.

Sperm whales do spend extended periods at the surface and it may be reasonable to assume that for a sperm whale at the surface as the aircraft flies overhead detection is certain if the whale is on or close to the trackline. For each sperm whale, the probability that it is at the surface at the moment the aircraft passes overhead is equal to the proportion of time it spends at the surface. Let K represent the population value for that proportion. Then, if the sampling procedure is to count and estimate the distance, $y$, to each whale visible at the surface as it passes abeam, $\hat{\mathbf{K}} / \hat{\mathrm{f}}(0)$ provides an intuitively reasonable estimate for esw. By restricting the data to whales at the surface as they pass abeam, the expected value of the estimator is made independent of the sighting condition and observers, providing the assumption "whales which are at the surface and on the trackline are certain to be counted' holds. Whether that assumption does hold can be checked by using independent observers.

For most other species the proportion of time spent at the surface is extremely small and the method suggested above is not useful. However, it is feasible to estimate the number of surfacings occurring within a defined region ahead of the platform and combine this with measurements of the mean rate of surfacing for the target species in order to estimate abundance. This technique has come to be


Fig. 1. The region of sea surface, relative to the position of aircraft and trackline, scanned by observers during an aerial survey for fin whales near Iceland. It was assumed that all blows on or just beyond the line segment $S$ were certain to be seen. The broken lines indicate the positions of cut points between successive distance intervals. The figure also shows a whale blowing at distance f from line $S$, and distance $r+r_{0}$ from the aircraft, and the shaded region shows the corresponding value of the function $\mathrm{a}(\mathrm{r})$ (the region within distance $\mathrm{r}_{0}$ of the aircraft and angle $7^{\circ}$ from the trackline was not visible to the observers). (From Hiby et al., 1984.)
known as 'cue counting' and has been used in aerial surveys for fin whales near Iceland (Hiby, Martin and Fairfield, 1984), on the IDCR minke whale assessment cruises (Hiby and Ward, 1986a; Hiby and Ward, 1986b; Ward and Hiby, 1987) and on the North Atlantic Sightings Survey in 1987 (Hiby, Ward and Lovell, 1989). In order to define the moment of surfacing in an objective way the sighting of the blow has normally been used. However, any convenient definition can be used. For example, it may be more advantageous in aerial survey to use the start of the next dive. The basic data required are the number of cues observed within the defined region and the distance to each cue as it occurred. The defined region is a sector of known angle delineated, for example, by sighting bars or lines placed within the field of view of the observer(s); it has no 'hard' leading edge (Fig. 1). An estimate of effective search area (esa) is required from the data. This is the analogue of the esw in line transect and is defined as

$$
\int_{0}^{A} g(a) d a
$$

where $g(a)$ is the conditional probability of counting a cue occurring within the sector at distance $r$ from the platform, such that an arc of radius $r$ encompasses a region of area a within the sector (see Fig. 1). In practice it may be better to measure distances and areas from a nearby arc at radius $r_{o}$ rather than from the platform itself. Intuitively, esa represents the ability of the observers on the platform to detect and record cues within the sector.

The derivation of the estimate of esa is analogous to the line transect case:

Let w represent a convenient truncation radius within which all or most of the detected cues occur. Let A represent the area encompassed within the sector within an arc of radius $w$. Assuming a cue occurring within $w$ is equally likely to occur anywhere in area $A$, that is, the pdf for the position of the cue is uniform at $1 / \mathrm{A}$ then the probability the cue is seen is

$$
1 / A \int_{0}^{A} g(a) d a
$$

The probability the cue is seen $a t a<a *$ is

$$
1 / A \int_{0}^{a^{*}} g(a) d a
$$

and the probability the cue is at $a<a^{*}$ given it is seen is

$$
\frac{1 / A \int_{0}^{a^{*}} g(a) d a}{1 / A \int_{0}^{A} g(a) d a}=\frac{\int_{0}^{a^{*}} g(a) d a}{\text { esa }}
$$

Thus the pdf for sighting position with respect to the random variable a is
and

$$
f(a)=g(a) / e s a
$$

$$
\mathrm{esa}=\mathrm{g}(\mathrm{a}) / \mathrm{f}(\mathrm{a})
$$

In particular

$$
\mathrm{esa}=\mathrm{g}(0) / \mathrm{f}(0)
$$

Estimation of esa is thus analogous to estimation of esw from line transect data. $f(a)$ is estimated by curve fitting to the frequency distribution of distance (expressed in terms
of area $a$ ) to cues, and then evaluated at $a=0 . g(0)$ may be assumed to equal 1 , or independent observers may be used and the proportion of duplicate cues used to estimate $g(0)$ (see Section 8.3 ). Note that $g(0)$ is the probability of detecting a cue occurring within the sector at $a=0$, not the probability of detecting a whale or pod, so the problem of whales being invisible on dives as the platform passes by does not effect the estimate of $g(0)$ in the cue counting case.

Detection probabilities are clearly not independent for different cues (because a number of cues may be recorded on each pod surfacing and also because, depending on platform speed and surfacing rate, more than one cue may be recorded for each whale) so variance estimates are based on jackknife or bootstrap methods.

An estimate of local density, if required, is

$$
\hat{\mathrm{D}}=\mathrm{n} / \mathrm{BTesa}
$$

where $n$ is the number of cues counted during $T$ hours on effort and $B$ is the mean surfacing rate - surfacing per whale per hour - for the whales in that area.

Estimation of surfacing rates and the question of how the esa is incorporated into an estimate of abundance is considered in Section 8.

The cue-counting estimator is highly sensitive to whale reaction to the platform because of the assumption that the probability distribution for position of cue occurrence within the observed sector is uniform. This presents a serious problem for cue-counting sampling from shipboard surveys which requires further investigation.

### 2.6 Shipboard acoustic survey

Any of the techniques applicable to sightings are potentially available for acoustic surveys also. Detection of cetaceans by recording and identification of their vocalisations may provide greater effective search range than the use of visual cues (Thomas, Fisher, Ferm and Holt, 1986; Whitehead and Gordon, 1986). Those acoustic survey exercises carried out to date have not been designed to estimate effective search range; however, there does not seem to be any fundamental problem in deploying towed hydrophone arrays which permit the location of each source of vocalisation detected, as in static arrays (see Section 2.8). Sperm whales would appear to be particularly suitable candidates for this type of approach because of their unique vocalisation patterns. One possibility would be the use of an onboard microprocessor to obtain an estimate of location for each 'click' detected in real time, using the pattern of arrival time differences from the array. A series of clicks received from a whale would then identify its track relative to the vessel against the scatter of locations due to background noise. The number of those tracks crossing the abeam line and their distance from the trackline would be noted and subjected to the same analysis as described at the start of the last section - K representing the proportion of time spent in vocalisation. Given sufficient development, such equipment might also be suitable for use from platforms of opportunity.

Another possibility requiring less technological development would be to modify the spot sampling approach of Alling, Gordon, Rice and Whitehead (1983) to calculate the average effective range of the hydrophones used. One way to do this would be to record the whales detected at each station using a single $360^{\circ}$ sweep of a hydrophone which has a narrow angle of reception in the horizontal plane. This allows not only the number of
whales heard but also the compass bearing to cach whale to be recorded. On moving to a new station, the correlation in the pattern of bearings with that obtained from the previous station is related to the range of the hydrophone as compared to the distance between stations. Repeated sweeps at the same station are used to quantify the rate of decay in correlation resulting from movement of the whales and the commencement and cessation of clicking periods. The potential of this approach is being investigated.

It is worth noting here that there is a potential for estimating sperm whale size from their clicks (Norris and Harvey, 1979; Gordon, 1987).

### 2.7 Land-based visual survey

Although land-based surveys for whales are frequently referred to as censuses, it is more convenient and correct to consider them as sampling surveys within the same framework as line (and strip) transects. There are some differences between line transect and land-based survey methodology. An obvious one is that in a land-based survey the observer is stationary as the whales move past while in a line transect survey the observer moves through the whales which usually are assumed to be stationary. A more important difference is that in land-based surveys it cannot be assumed that the offshore distribution of whales is uniform, as is the case for line transects. In the absence of additional data, therefore, land-based surveys are more analogous to strip transects where it is assumed that all whales within a given perpendicular distance of the survey platform are certain to be seen and their actual distribution within the strip is irrelevant. It is because of this strip transect assumption that land-based surveys are often thought of as censuses. If information on the offshore distribution of whales is also available, it is quite straightforward to fit land-based surveys into the framework of line transects (see e.g. Hammond, 1984a).

If all whales passing within a given distance of the land are certain to be seen, no whales pass outside this strip and no whales pass by before or after the survey, then we can state that a land-based survey is a census. However, there are several ways in which these assumptions could be violated. Firstly, whales may pass the survey point before or after the survey period. These whales have a zero probability of being sampled and are therefore excluded in any population count. In land-based surveys for gray whales at Monterey, California (Reilly, Rice and Wolman, 1983) and at Unimak Pass, Alaska (Rugh, 1984) the number of whales missed before and after the surveys was estimated by fitting a predictive model to the daily counts and extrapolating it to include the tails of the distribution. Secondly, whales may pass outside the visual strip. These whales will also not be included in the population count. Additional data are required to estimate the proportion of the population which passes the survey point out of visual range (see Section 3.5). Thirdly, whales may be missed within the visual strip because the probability of detecting them is a decreasing function of distance from land. In this case a land-based survey is analogous to a one-sided line transect survey. The number of whales missed within the strip can be estimated if additional data are available, e.g. for bowhead whales surveyed at Point Barrow, Alaska. (Zeh, Ko, Krogman and Sonntag, 1986a, b). Fourthly, poor visibility may result in whales passing within range being missed. In practice it is necessary to define conditions of unacceptable visibility. Those periods can
then be corrected for by interpolation between periods of acceptable visibility as was done for the gray whale surveys at Unimak Pass and at Newport, Oregon (Herzing and Mate, 1984) and the bowhead whale surveys at Point Barrow (Zeh et al., 1986a, b), or by using a predictive model as for the gray whale surveys at Monterey. Finally, whales can only be counted during daylight hours so the number passing during the night must be estimated. Experimental data have been collected for gray whales at Monterey (see Section 3.5) to allow this. The bowhead surveys at Point Barrow take place at a time (mid-April to mid-June) when there is continuous light, so no correction is needed.

If these five factors can be addressed, land-based visual surveys are a cost-effective way of estimating the abundance of populations which migrate close to land (or ice) at a specific time of year.

### 2.8 Land-based acoustic survey

In an acoustic survey, it is the sounds made by the whales which are sampled rather than the whales themselves. Clearly, some whales may be heard more than once and others not at all. In order to convert sounds into animals it is necessary either to calibrate the number of sounds with independent data on sounds made per whale per unit time or to locate where each sound was made and join them up to give a series of 'whale-tracks' and hence a minimum number of whales. In the bowhead surveys at Point Barrow, efforts to correlate whale sounds with visual sightings have proved unsuccessful (Ko, Zeh, Clark, Ellison, Krogman and Sonntag, 1986). However, an algorithm has been developed (Sonntag, Ellison, Clark, Corbit and Krogman, 1986) which makes a maximum number of connections between the located sounds to produce whale 'tracks' given a range of allowed whale swimming speeds and directions. This produces a minimum count of whales from the sounds. Sonntag, Ellison and Corbit (1988) have investigated the sensitivity of the whale tracking algorithm to variation in its input parameters. For the bowhead whale, a minimum count of animals located acoustically out of the visual range of shore-based observers can be added on to a minimum visual count to give a better estimate of abundance.

Land-based acoustic surveys require sophisticated equipment for data collection and analysis and are only appropriate where the expertise exists to develop and deploy such equipment.
2.9 Combination of land-based visual and acoustic surveys Gentleman and Zeh (1987) and Zeh, Turet, Gentleman and Raftery (1988) have presented a method of combining data from the ice-based visual and acoustic surveys of bowhead whales from Point Barrow, Alaska. These authors use visual sightings (see Section 2.7) and acoustic whale 'tracks' (see Section 2.8) as the first and second samples in a simple two-sample mark-recapture model. The whales detected both visually and acoustically are then the recaptures. Zeh et al. (1988) used this method to estimate the number of bowheads passing Point Barrow within 3 km of the ice edge from the 1985 data. Whales detected acoustically more than 3 km from the ice edge were added to this estimate as before. This method represents a considerable improvement in the estimation of bowhead population size. Further analysis of more recent data and the development of an empirical approach to the estimation of confidence intervals for these estimates
(Raftery, Turet and Zeh, 1988) promise to increase the quality of this procedure and, therefore, the accuracy and precision of the estimates.

## 3. DATA COLLECTION

In this Section we do not attempt to present a comprehensive review of all aspects of data collection for all types of survey. Rather we have selected those topics which we consider would benefit from discussion of their associated problems.

### 3.1 Estimating the position of each sighting relative to the platform

Each survey technique (other than strip transect) requires an estimate of the radial or perpendicular distance to each sighting. In line transect sampling perpendicular distance, $y$, to sightings may be estimated as they come abeam or be calculated from estimates of the radial distance, $r$, and angle from the trackline, $\theta$. In the hazard rate analysis of line transect data (Section 2.4) the $r$ and $\theta$ values themselves are used in the analysis.

For unbiased estimation of abundance it is necessary that these estimates of distance and angle are unbiased, at least for small $y$ (or $r$ in the cue counting case) (see Section 4.1.2). The effects of variance in the estimates have not been considered formally because the models for estimating effective search range have not incorporated the error structure in the distance and angle estimates. However it is reasonable to suppose that the errors should be kept as small as possible.

Shipboard surveys have often relied upon the observer's innate ability to estimate distances at sea and, in the case of the IDCR surveys, calibration factors derived from the results of distance estimation experiments have been applied (e.g. Butterworth, Best and Hembree, 1984). However, observers vary widely in their ability to estimate distance, the estimation experiments may not be representative of the problem of estimating distances to sightings and severe problems in the analysis of survey data based on unaided distance estimates have been encountered (Whitehead, 1982).

The only type of instrumentation that has been used routinely on shipboard surveys is a scale to estimate the angle of declination from the horizon to the sighting. The angle of declination to the sighting is simply related to distance given the height of the observer above sea level and the curvature of the earth. The scale may be hand-held or incorporated as a graticule into the binoculars used for searching, as in the IDCR cruises. It probably serves more as a way of continually reinforcing the observer's ability to judge distance than as a direct measurement scale, because the potential for aligning a scale with a fleeting glimpse of a whale from a moving ship is very limited. Experimental results suggest that the use of scales have lasting influence on distance estimation even when they are no longer deployed (Thompson and Hiby, 1985). The measurement of declination method has the very desirable property that the relative error decreases with decreasing distance (see Section 4.1.2). Observers should, however, be aware of the highly non-linear relationship between declination angle and distance, otherwise linear interpolation between marked distances results in upward bias in distance estimates (Thompson and Hiby, 1985).

Unaided estimation of angle from the trackline (in line transect sampling) can also cause serious problems in data analysis. This is mainly due to the tendency to round angles to convenient values to an unknown and variable degree. Angles rounded to zero cause a spike in the distribution of perpendicular distance $y(e s t i m a t e d ~ a s ~ r \sin \theta)$ at $y=0$ which precludes reliable estimation of $f(0)$ (see Section 5). Rounding to zero may be particularly prevalent because of the tendency of the vessel to yaw around the trackline. As a result a sighting a few degrees to port or starboard of the trackline may come to lie directly ahead of the vessel within a few seconds and then be recorded as 'ahead'.

Various 'smearing' techniques (Section 4.1.2) have been used to overcome this problem in the data (Butterworth, 1982a; Hammond, 1984b), but the estimator of $f(0)$ is sensitive to the degree of smearing used. A recent development by Buckland and Anganuzzi (1988) should help to determine the degree of smearing required (see Section 5.1); however, it is preferable to try to reduce or eliminate the problem of rounding by using an 'angle board' for angle estimation (Joyce, Nakanishi, Hata and Pastene, 1985). In the IDCR surveys the binoculars used for searching are mounted on a supporting pole fitted with a pointer which rotates across a fixed angular scale. The observer centres the sighting in the binoculars field of view and then reads off the angle from the scale. A similar pointer/scale device can be constructed for observers searching without binoculars. In either case the positioning of the pointer avoids the tendency to round angles; the only rounding is to the resolution of the scale, which is known. The error (as opposed to rounding) induced by the yaw of the ship can be eliminated by signalling to the bridge to take a simultaneous recording of the heading of the ship at the moment the pointer is aligned with the sighting. The use of an angle board on the IDCR cruises since 1984 appears to have eliminated the spike apparent in the earlier data sets (Buckland, 1987). Angle boards have also been used successfully on the North Atlantic Sightings Survey in 1987.

In aerial surveys, perpendicular distance is normally measured to sightings as they come abeam, either by measuring the angle of declination using markings on the windscreen (Scott and Gilbert, 1982; Hiby, Martin and Fairfield, 1984) or measuring the angle from the horizontal, using an inclinometer. Because of the greater height of the observer above sea level these measurements are far more accurate from aircraft than from ships. Another technique, used during aerial surveys of dolphins in the eastern tropical Pacific, is to calculate the distance moved by the aircraft to close with a sighting using a navigational computer.
For the cue counting technique, radial distances to cues are required. Because of the speed of the aircraft, any delay in obtaining inclinometer measurements causes a downward bias in the corresponding distance estimates, particularly for cues seen at small angles to the trackline. One way to overcome the problem is to record data onto audio tape, indicating both the moment the cue occurred and the moment the inclinometer reading was obtained. On transcription of the data the delay time can be measured and a correction applied, based on the speed of the aircraft, the perpendicular distance at which the whale passed abeam of the aircraft and the angle of drift of the aircraft. The procedure is greatly simplified if data are recorded onto one track of a stereo tape recorder and the second track is used to record a continuously updated time
signal, which is displayed on replay. The time of each recorded event is then transcribed and the distance estimates and corrections required are derived at the stage of data analysis. Recording equipment of this type was used successfully on the 1986/87 IDCR cruise (Ward, pers. comm.) and in the North Atlantic Sighting Survey in 1987 (e.g. Donovan and Gunnlaugsson, 1989).

### 3.2 Sighting conditions

A number of researchers have attempted to formalise the effect of environmental conditions on the sightability of different cetaceans, e.g. Clarke (1982). The Cetacean and Turtle Assessment Program (Scott and Gilbert, 1982) placed great emphasis on the collection and analysis of environmental data, and recording of such data forms part of all cetacean surveys. However, the objectives of collecting such data have not received much consideration, and neither has the selection of the types of data to be collected. If the objective is to derive calibration factors to allow comparison of different surveys, this may not be achievable. Even if an environmental factor can be demonstrated to have a significant effect on the estimate of abundance, e.g. Holt (1984), the size of that effect may be very difficult to quantify and may well depend upon the observers used on that survey. Surveys must be based on a technique in which the expected value of the abundance estimate is independent of the sighting conditions (the sighting rate and effective search range will not, of course, be independent of the sighting conditions). Environmental data do, however, have an important role in defining minimum conditions for the chosen technique to provide reliable estimation of abundance and to provide an objective criterion for rejecting data which were collected under unacceptable conditions.

Which types of data are worth collecting? It would seem worthwhile to consider the physical structure of the sighting process and obtain specialist advice concerning the factors determining the intrinsic and apparent levels of contrast between the various cues and the various types of background (i.e. sea surface from the air, sea surface or sky from a ship). For example, what are the environmental factors affecting the density of a blow? Is there some objective way (analogous to the Beaufort scale) to describe the colour of the sea surface, for example, a reference chart? Such information would facilitate the use of environmental data in analysis. (See Anon., 1987, for further discussion of this topic.)

### 3.3 Independent observers

The use of two sets of observers to estimate $g(0)$ in line transect sampling is mentioned in Sections 2.2 and 2.3. This section describes requirements of this procedure in terms of data collection. Section 4.2 discusses the analysis of such data.

In the basic method, each set of observers is unaware of those sightings by the other set. This precludes the use of closing mode sampling procedures (Section 2.3). Sightings may be allocated to duplicate or non-duplicate classes by a third observer who is informed of the sightings by both sets as they occur, or the allocation may take place at a later stage (Butterworth, Best and Hembree, 1984). The allocation is based on sighting position, time when abeam, and (in the case of real-time allocation only) the degree of synchrony in the timing of blows from the sightings observed by each set. In the case of cue counting sampling,
cues are allocated to the duplicate category during data analysis on the basis of time of occurrence; for this it is necessary that data are recorded onto stereo tape recorders as described in Section 3.1.

If it is necessary for the platform to break off the trackline to confirm the species or pod size of a sighting the method described above breaks down. However, if the vessel turns onto sightings made by one set of observers only, a revised analysis by Cooke (Section 4.2) based on a hazard rate approach, permits estimation of $g(0)$ for each set. The data required are the number of sightings made by the set whose sightings are closed with, the number made only by the other set, and the number made by the other set which are subsequently detected by the first set. The method assumes that the hazard rate functions for the two sets of observers are proportional. To ensure that this is the case the roles of the two sets need to be exchanged regularly, that is, the vessel would close on sightings by the first and second sets in alternate periods.

Problems in analysis arise if the observers allocate effort to markedly different regions within the same $y$-intervals (or, for cue counting sampling, different regions of the same r-intervals). For example, one set may search further ahead than the other, or there may be unequal allocation of effort to port and starboard. It may therefore be worthwhile trying to make some measurements of allocation of effort. Certainly, all available information on the position of sightings should be recorded.

### 3.4 Survey platforms

IWC (1982, pp.547-9) reported a summary of inter-platform comparison exercises that had been carried out by that time and further comparisons have been completed since then (e.g. Kraus, Gilbert and Prescott, 1984; Withrow, Rice and Wolman, 1983).
That report stated that 'general conclusions were difficult to draw'. This is not surprising given the lack of an established basis for comparison. For example, when comparing 'variability' of abundance estimates, should the comparison be for estimates obtained over equal transect length, equal times on effort or for equal cost of operation for the platform? Such comparisons are relevant to the choice of platform for a given survey, but only if they can be related to the constraints applying in the actual situation.

When contemplating carrying out a survey there may be no need to make a choice of platform in that the survey may be conditional on the use of a given platform. If a choice is to be made, the constraints applying will likely be cost, the size and 'remoteness' of the region to be surveyed and the time available for the survey. The general types of platform available (apart from land-based surveys) are ships and aircraft. Yachts represent a special category, unsuitable for sightings but suitable for certain types of acoustic survey. Given the early stage of development of shipboard acoustic surveys they are not considered here.
In coastal areas, say no more than one to two hundred miles from the nearest landing strip, light aircraft are many times cheaper to operate per transect mile on effort. Although hourly rates work out higher for aircraft, the higher speed and lower proportion of hours off-effort combine to give the lower cost per transect mile. For more remote areas, involving long transit flights, the proportion of hours flown off-effort increases and aircraft may become uneconomic. The use of a helicopter flying transects from a
ship transiting a remote area may be cost-effective; surveys conducted in Alaska and Antarctica indicate that such an arrangement is feasible (G. Joyce, pers. comm.).

The effective search range will normally be many times lower from an aircraft than from a suitable vessel. Thus it is not possible to say, in general, which type of platform will give more sightings per unit cost. However, sightings from aircraft are more effective in that the effort is more finely divided, with much shorter inter-transect distances. As a result, aerial survey data avoid the component of variance due to aggregation of whales at scales greater than the mean inter-transect distance. The use of aircraft also avoids the effects of the various types of whale movement, and eliminates problems of surveying irregular coastlines and moving ice edges.

We suggest that these considerations favour the use of aircraft in areas where transit times are not too high, as long as the sampling technique is one capable of providing unbiased estimation. For large whales, this implies the use of cue counting methods based on the measurement of the proportion of time at the surface or surfacing rates (see Section 2.5 and Section 8). The use of independent observers is also necessary, either to check the $g(0)=1$ assumption or to estimate the ratio $g(0) / f(0)$, and also to quantify the degree of error in estimates of distance to cues. Where a shipboard survey is used, either because the area is unsuitable for an aerial survey or because of the availability of a vessel, the use of independent observers as an integral part of the sampling procedures in highly desirable. Whether the cue counting approach to sampling should be used on ships is less clear. The sensitivity of the cue counting estimator to vessel reaction is, potentially, a serious problem. Other problems remain for line transect sampling, for example, the estimation of mean pod size appropriate to the estimate of $g(0) / f(0)$, and the estimation of $g(0) / f(0)$ for species like sperm whales which undertake very long dives. We suggest that, at this stage, data collected should permit alternative analyses.

### 3.5 Land-based visual surveys

Compared to collecting data from aerial or shipboard platforms, data collection in a land-based visual survey is relatively straightforward. In its simplest form, whales are counted as they pass the survey point. In order to investigate the distribution away from the shore of detected whales it is necessary to measure their distance from the survey point. This can be done with inclinometers or, more accurately, by using theodolites as is done on the bowhead surveys. Measurement of the bearing of the sighting from the survey point allows the position of the whale to be fixed. This facilitates the collection of data from which the proportion of whales missed within the visual strip can be calculated and provides data which could allow alternative analyses to be performed (e.g. line transect - Hammond, 1984a).

Estimating the number of whales missed within the visual strip requires at least two independent observers as described in Sections 2.3, 3.3 and 3.4 for the line transect and cue counting methods. This is achieved at the annual bowhead surveys by setting up two counting locations, known as perches, on the shore-fast ice, close enough together so that it can be assumed that the probability of detecting each approaching whale is the same at each perch. The proportion of whales missed is estimated using the so-called 'removal' method (Seber, 1982, p.318) by treating the whales seen at each perch as independent
samples. The $n_{1}$ whales seen at the first perch are 'removed' from the population by notifying the second perch that they have already been seen and $n_{2}$ of the whales missed by the first perch are then counted at the second perch. An estimate of the proportion of whales missed is then simply $n_{2} / n_{1}$. Double-counting experiments at Monterey have shown that a similar correction may also be necessary for gray whales (Cooke, 1986a).

Whales passing outside the visual strip can only be accounted for if additional data are collected such as from aerial surveys. This has been done in the bowhead surveys and gray whale surveys at Monterey and Newport.

Data collected for gray whales have been used to investigate whether or not the night-time migration rate is different from the daytime migration rate. Two experiments at Unimak Pass using night goggles at dusk showed a slight slowing to $73 \%$ of the daytime migration rate in one case and no differences in the other (Rugh, 1984). More recently, a telemetry study of eighteen individuals during the 1985/86 Monterey census showed no difference between daytime and night-time migration rates (Swartz, Jones, Goodyear, Withrow and Miller, 1987). These results suggest that daytime counts can simply be multiplied up proportionally to estimate the daily numbers of gray whales passing a survey point.

### 3.6 Acoustic surveys

Sophisticated sonic equipment may be needed to pick up the whale sounds, record or monitor them, and to locate the sounds in space if required. Clearly, such surveys are not undertaken without a great deal of expertise and investment. Much of the pioneering work in shore-based acoustic surveys has been done as part of the bowhead whale surveys at Point Barrow, Alaska. Readers interested in the details of land-based acoustic data collection are referred to Ellison, Clarke and Beeman (1985), and in shipboard data collection to Thomas et al., (1986).

## 4. DATA ANALYSIS

As in Section 3, we have not tried to cover all aspects of data analysis in this Section but have chosen to concentrate on certain topics which have caused problems in analysis of cetacean survey data.

### 4.1 Estimation of $f(0)$

It was suggested in Sections 2.3 and 2.5 that the general problem in estimating an effective search range from line transect or cue counting sample data was that of estimating $g(0) / f(0)$. In some cases, however, the assumption $g(0)=1$ will be justified and in many cases data for estimation of $g(0) / f(0)$ will not be available, and it is therefore also relevant to consider the properties of estimators for $f(0)$ in isolation from the problem of estimating $g(0)$. Section 5 deals with the choice of a functional form for $f(y)$. This section considers two related questions: the presence of a 'shoulder' in $g(y)$ and the effects of errors in distance estimation and species identification. There is inevitably some overlap of material in Sections 4 and 5.

### 4.1.1 Should the data exhibit a 'shoulder'?

Burnham et al. (1980) suggested that any model for $g(y)$ should have a 'shoulder', i.e. that $\mathrm{dg} / \mathrm{dy}$ should be zero at $y=0$; this condition they called the 'shape criterion'. Given random placement of transects the distribution of $y$ to stationary targets will be uniform, at least on the scale of
the visible range, $w$, so that a shoulder in $g(y)$ should be apparent in $f(y)$. However, many data sets from cetacean surveys do not exhibit a shoulder, at least at the resolution available, and tend to be 'spiked' at $y=0$. Data sets which are spiked in this way preclude reliable estimation of $f(0)$ (although not necessarily of $g(0) / f(0)$, see Section 4.2) and there has therefore been considerable discussion concerning the possible reasons why data may fail to show a shoulder, such as bias in estimation of distance and angle to sightings, concentration of observer effort along the trackline, whale response to vessel and the effect of random whale movements. As a basis for such discussion it is necessary to consider whether, or to what degree a shoulder is to be expected given accurate measurement of $y$ and no whale movement.

Koopman (1980) and Butterworth (1982b) showed that for a stationary target at perpendicular distance $y$ from the trackline, the probability of detection

$$
g(y)=1-\exp \left\{-(1 / V) \int_{0}^{\infty} h[r(x, y), \theta(x, y)] d x\right\}
$$

assuming zero search effort to the rear of the abeam line. The function $h(r, \theta)$ is the hazard rate giving, as $h(r, \theta) d t$, the probability that a target at position $\mathrm{r}, \theta$ will be detected during the interval $\mathrm{dt} . \mathrm{V}$ is the speed of the vessel. This formulation has been used to investigate the implications of alternative assumptions about the hazard rate for the shape of $g(y)$ (Butterworth, 1982b) and to derive the 'hazard rate model' for fitting $\mathrm{f}(\mathrm{y})$ (Buckland, 1985).

Differentiating with respect to $y$ we obtain:

$$
g^{\prime}(y)=\exp \left[-(1 / V) \int_{0}^{\infty} h(r, \theta) d x\right] \cdot(1 / V) \int_{0}^{\infty}\left[\frac{\delta h}{\delta r} \frac{d r}{d y}+\frac{\delta h}{\delta \theta} \frac{d \theta}{d y}\right] d x
$$

The derivative $\mathrm{dr} / \mathrm{dy}$ is zero at $\mathrm{y}=0$ so for $\mathrm{g}^{\prime}(0)=0$ it is sufficient that at $\mathrm{y}=0, \delta \mathrm{~h} / \delta \mathrm{r}$ is finite and $\delta \mathrm{h} / \delta \theta=0$. At small $\theta$, the dependence of the hazard rate on $\theta$ is dominated by allocation of effort with angle by the observers, and although measurements suggest observers do concentrate their effort ahead (Doi et al., 1982; Thompson and Hiby, 1985; Ward, Hiby and Thompson, 1986; Kasamatsu and Kishino, 1986) the allocation of effort is not spiked at $\theta=0$. If the partial derivative $\delta h / \delta r$ is not finite at $y=0$, as for example in the second case considered by Butterworth (1982b), then $g(y)$ may by spiked but only if the distribution of $x$ distances to sightings along the trackline is also spiked, i.e. we would need a high concentration of sightings very close to the vessel. The burden of this argument is that if the data are spiked with respect to $y$ but not with respect to $x$, even at small $y$, and if, furthermore, the allocation of effort with angle is not spiked at zero, then the spike in $y$ is either an artifact of the method used to estimate $y$ or the result of whale movement. The situation poses considerable problems for analysis - smearing the data (Section 4.1.2) or the use of a hazard rate analysis are possible approaches. Section 5 considers further the question of 'spiked' data.

### 4.1.2 Errors in distance estimation

Consideration of the effects of error in distance estimation is complicated by the fact that although the definition of esw is in terms of $g(y)$, where $y$ is a continuous random variable, the data are in terms of distance estimates which are realisations of a discrete random variable. For example, although it is convenient to talk loosely about the desirability of unbiased distance estimation, that concept is not really sufficient. If estimates of distance were given to
the nearest half mile we would presumably wish a true distance of 1.1 n .miles to be estimated at 1 n .miles, whereas for estimation of that true distance to be unbiased it would be necessary to have some estimates at 1.5 n.miles.

If the estimates obtained were simply a distorted version, $y^{\prime}$, of the true distances, $y$, then estimation of esw would be unbiased providing the functional form chosen for $\mathrm{f}(\mathrm{y})$ was 'model robust' in the sense of Burnham et al. (1980, p.44) and the distortion approached zero as $y$ approached zero. This is because the probability density function for $y^{\prime}$ equals

$$
\mathrm{f}\left[\mathrm{y}\left(\mathrm{y}^{\prime}\right)\right] \cdot\left(\mathrm{dy} / \mathrm{dy} y^{\prime}\right)
$$

so that if $d y / d y^{\prime}=1$ at $y=0$, then $y$ and $y^{\prime}$ have the same probability density at zero. This suggests that distance estimates need only be 'unbiased' at short distances. That conclusion does not apply if the functional form chosen for $f(y)$ is not model robust. For example, for the negative exponential, the esw estimate depends on distance estimators only via the mean, $\overline{\mathrm{y}}$, which is affected by bias at all distances.
In practice, $f(0)$ is not estimated from a sample of $y^{\prime}$ values but from a set of $m$ class frequencies $n_{i}$, i.e. the number of sightings estimated to be within each of $m$ distance intervals. The frequencies $n_{i}$ are regarded as one realisation of a multinomial distribution with probabilities $f_{i}$, equal to the integrals of the density $f(y)$ over the corresponding distance intervals. The likelihood

$$
\prod_{i=1}^{m} f_{i}^{n_{i}}
$$

is maximised with respect to the parameters of the function chosen to represent $f(y)$ and then $f(y)$ is evaluated at $y=0$. Errors in distance estimation thus affect $f(0)$ via the frequencies $n_{i}$, misclassifications among the intervals closest to the trackline having the greatest effect.

The problem of misclassification may be severe for line transect analysis because sightings at long radial distances can have small y values. Furthermore, rounding to certain values of $r$ and $\theta$ gives rise to misclassification of $y$. For example, sightings seen between 1.3 n .miles and 1.7 n.miles may be rounded to 1.5 n.miles. Such sightings occurring and correctly estimated at $10^{\circ}$ from the trackline would then be assigned a $y$ value of 0.262 n.miles and be allocated to the 0.25 to 0.3 n.miles $y$-interval. However, sightings between 1.3 n.miles and 1.44 n .miles from the vessel should actually have been allocated to the 0.2 n .miles to 0.25 n .miles y -interval. 'Smearing' is designed to reduce the effect of this type of misclassification (Butterworth, 1982a). Various methods of smearing were considered by Hammond (1984b) in relation to data from the IDCR minke whale assessment cruises. The basic idea is to determine for each sighting, the possible range of $\theta$ and $r$ values within which it could have occurred and then to calculate the probability of it having occurred in each of the $y$-intervals having some overlap with the range of $r$, values. Recent work by Buckland and Anganuzzi (1988) concerns estimation, from the data, of the degree of rounding in $\theta$ and $\mathbf{r}$ occurring in the data.

### 4.1.3 Species misidentification

Misidentification of species should not cause serious problems if it is unlikely to occur for targets seen close to the trackline (or in the cue counting approach, for cues seen close to the platform). It is intuitively obvious that
failing to identify whales of the target species is similar to failure to detect them and should therefore not lead to biased estimation of abundance. It is less obvious that erroneously including whales of other species will not lead to bias; however, the following example shows that although the estimate of esw is increased, this is compensated for by a proportional increase in the expected number of targets counted, so long as the probabilities of misidentification reduce to zero at $\mathrm{y}=0$ (or $\mathrm{r}=0$ ).

Suppose that along L miles of transect, $\mathrm{N}_{1}$ pods of target species 1 and $\mathrm{N}_{2}$ pods of 'nuisance' species 2 are encountered within w miles of the trackline. Let $P_{11}(y)$ be the probability that a pod of species 1 is (correctly) identified as species 1 , given it is at distance $y$ from the trackline, and $P_{21}(y)$ be the probability that a pod of species 2 is (erroneously) identified as species 1 . Let $P_{11}(0)=1$ and $P_{21}(0)=0$. Let $g(y)$ and $g(y)$ be the detection functions for species 1 and 2 , respectively.

The probability that a pod encountered within distance w is detected is then

$$
\frac{N_{1}}{N_{1}+N_{2}} \cdot \frac{1}{w} \int_{0}^{w} g_{1}(y) P_{11}(y) d y+\frac{N_{2}}{N_{1}+N_{2}} \cdot \frac{1}{w} \int_{0}^{w} g_{2}(y) P_{21}(y) d y
$$

The probability the pod is detected within distance $y^{*}$ is

$$
\frac{N_{1}}{N_{1}+N_{2}} \cdot \frac{1}{w} \int_{0}^{y^{*}} g_{1}(y) P_{11}(y) d y+\frac{N_{2}}{N_{1}+N_{2}} \cdot \frac{1}{w} \int_{0}^{y^{*}} g_{2}(y) P_{21}(y) d y
$$

Therefore, the probability it is within distance $\mathrm{y}^{*}$ given it is detected is

$$
\frac{N_{1} \int_{0}^{y^{*}} g_{1}(y) P_{11}(y) d y+N_{2} \int_{0}^{y^{*}} g_{2}(y) P_{21}(y) d y}{N_{1} \int_{0}^{w} g_{1}(y) P_{11}(y) d y+N_{2} \int_{0}^{y^{*}} g_{2}(y) P_{21}(y) d y}
$$

The pdf for distance to sightings is

$$
f(y)=\frac{N_{1} g_{1}(y) P_{11}(y)+N_{2} g_{2}(y) P_{21}(y) d y}{N_{1} \int_{0}^{w} g_{1}(y) P_{11}(y)+N_{2} \int_{0}^{w} g_{2}(y) P_{21}(y) d y}
$$

and

$$
f(0)=N_{1} g_{1}(0) /\left\{N_{1} \int_{0}^{w} g_{1}(y) P_{11}(y)+N_{2} \int_{0}^{w} g_{2}(y) P_{21}(y) d y\right\}
$$

Thus esw $=g_{1}(0) / f(0)$

$$
=\int_{0}^{w} g_{1}(y) P_{11}(y) d y+\left(N_{2} / N_{1}\right) \int_{0}^{w} g_{2}(y) P_{21}(y) d y
$$

and is therefore increased by inclusion of some pods of species 2.
The expected number of targets counted, $n$, is

$$
N_{1}(1 / w) \int_{0}^{w} g_{1}(y) P_{11}(y) d y+N_{2}(1 / w) \int_{0}^{w} g_{2}(y) P_{21}(y) d y
$$

and is therefore increased by the same factor, i.e.

$$
1+\left(N_{2} / N_{1}\right) \int_{0}^{w} g_{2}(y) P_{21}(y) d y / \int_{0}^{w} g_{1}(y) P_{11}(y) d y
$$

### 4.2 Estimation of sighting efficiency from independent observer experiments

The first attempt to estimate $g(0)$ experimentally for whale surveys was by variable speed experiments conducted during the IDCR Southern Hemisphere minke whale assessment cruises (Butterworth, Best and Basson, 1982). However, the theoretical basis for such experiments was later found to be unsound (Cooke, 1985).

Recently, attention has focussed on the method of independent observers. The theory behind this method is that two sets of observers record sightings without knowledge of each other's sightings. From an analysis of the proportion of overlap in the two sets of observers' sightings records, an estimate of the number of whales missed by both observers can be obtained, on the assumption that the sighting of a whale or pod by one observer is probabilistically independent of the sightings of the same whale or pod by the other observer.

The first such experiments were conducted using observers on different ships steaming in parallel a fixed distance apart ('the Parallel Ship experiment') (Butterworth et al., 1982; Butterworth, Best and Hembree, 1984). However, the use of two separate vessels poses some extra logistical constraints, and makes the identification of duplicates (pods seen by both observers) somewhat harder since angular positions can differ greatly. Because of this, the experiments on recent IDCR cruises have had both sets of observers on the same vessel. This may facilitate duplicate identification, but requires a vessel whose layout enables two independent observer platforms which are audibly and visually isolated from each other.
The independence between observers cannot be maintained if the sightings are closed on. A modified method suitable for use during closing mode surveys is developed later.
The material in this section is taken largely from Cooke (1987a), where more detailed derivations and examples can be found.

### 4.2.1 Basic theory

Let $g_{A}(y), g_{B}(y)$ be the sightings probability of a pod a perpendicular distance $y$ from the trackline for the two sets of observers ( $\mathrm{A}, \mathrm{B}$ ) respectively. On the assumption of independence, the probability of being seen by both observers is $g_{A B}(y)=g_{A}(y) . g_{B}(y)$.

Assume the following notation:
$\mathrm{n}_{\mathrm{AB}}$ : number of pods sighted by both observers;
$\mathrm{n}_{\mathrm{A}}, \mathrm{n}_{\mathrm{B}}$ : number of pods sighted by observer A, B, respectively.

The probability that a whale is seen by observer A , given that it is seen by observer $B$, is:

$$
\begin{aligned}
& \mathrm{p}_{\mathrm{A} \mid \mathrm{B}}=\int_{0}^{\infty} \mathrm{g}_{\mathrm{A}}(\mathrm{y}) \mathrm{g}_{\mathrm{B}}(\mathrm{y}) \mathrm{dy} / \int_{0}^{\infty} \mathrm{g}_{\mathrm{B}}(\mathrm{y}) \mathrm{dy} \\
&=\mathrm{Fg}_{\mathrm{A}}(0) / \mathrm{f}_{\mathrm{A}}(0) \\
& \mathrm{F}=\int_{0}^{\infty} \mathrm{f}_{\mathrm{A}}(\mathrm{y}) \mathrm{f}_{\mathrm{B}}(\mathrm{y}) \mathrm{dy}
\end{aligned}
$$

where
and $f_{A}(y), f_{B}(y)$ are the pdfs of the perpendicular distance distribution of the two observers' sightings (so $\left.f_{A}(y) d y=f_{B}(y) d y=1\right)$.

Hence, conditional on $n_{B}$, an estimate of $g_{A}(0)$ is:

$$
\mathrm{g}_{\mathrm{A}}(0)=\mathrm{f}_{\mathrm{A}}(0) \mathrm{n}_{\mathrm{AB}} / \mathrm{Fn}_{\mathrm{B}}
$$

An equivalent estimate of $g_{B}(0)$ is also obtainable. An estimate of $g_{A+B}(0)$, i.e. the $g(0)$ for both sets together, is obtained by the relation:

$$
\mathrm{g}_{\mathrm{A}+\mathrm{B}}(0)=\mathrm{g}_{\mathrm{A}}(0)+\mathrm{g}_{\mathrm{B}}(0)-\mathrm{g}_{\mathrm{A}}(0) \mathrm{g}_{\mathrm{B}}(0)
$$

Conditional on $n_{B}$, the $C V$ of $n_{A+B}$ is $V\left[\left(n_{B}-n_{A B}\right) /\right.$ $\left(\mathrm{n}_{\mathrm{B}} \mathrm{n}_{\mathrm{AB}}\right)$ ] under certain independence assumptions. The CV of $\mathrm{g}_{\mathrm{A}}(0)$ will also have a component due to the CV of $f_{A}(0)$ and of $F$.

In practice, sightings at distances greater than some fixed truncation distance would be excluded from the estimation of $g(0)$ and $f(0)$.

### 4.2.2 Extensions to basic theory

## (a) Case where $f(0)$ is not estimable

Reliable estimation of $g(0)$ requires reliable estimation of $f(0)$. This will not be possible if $f(y)$ is sharply peaked at $y=0$. This is because the height of the peak can be varied substantially by changing the shape of the peak in the immediate neighbourhood of $\mathrm{y}=0$ without appreciably affecting the overall fit of $f(y)$ to the data, as illustrated in some of the examples in Section 5.

Such peaks are often observed in line transect data for whales, with some doubt as to whether they are real or artifacts (see Section 5).

However, the quantity of relevance for density estimation is not $g(0)$, but $g(0) / f(0)$. The formula for estimation of $g(0) / f(0)$ is:

$$
\left[\mathrm{g}_{\mathrm{A}}(0) / \mathrm{f}_{\mathrm{A}}(0)\right]_{\mathrm{EST}}=\mathrm{n}_{\mathrm{AB}} /\left(\mathrm{n}_{\mathrm{B}} \mathrm{~F}\right)
$$

This requires only the estimation of $F$, which, being an integral over the entire function, is far less sensitive to changes in the shape of $f$ near $y=0$. Two curves can yield almost the same F value, despite having greatly different $\mathrm{f}(0) \mathrm{s}$.

## (b) Case where $g(y)$ is not symmetrical

In conventional line transect theory, data from port and starboard sides of the trackline can be folded together; any asymmetry in the perpendicular distance distributions is not relevant.
However, for independent observer experiments, if both $g_{A}(y)$ and $g_{B}(y)$ are asymmetrical about $y=0$, then this will introduce some bias in the estimates of $g(0)$. This bias can be avoided by replacing the $f(0)$ and $F$ values in the estimator for $g(0)$ by a weighted combination of the two sides evaluated separately:

$$
\mathrm{F}=\left(\mathrm{n}_{-\mathrm{A}} \mathrm{n}_{-\mathrm{B}} \mathrm{~F}_{-}+\mathrm{n}_{+\mathrm{A}} \mathrm{n}_{+\mathrm{B}} \mathrm{~F}_{+}\right) /\left(\mathrm{n}_{-\mathrm{A}} \mathrm{n}_{-\mathrm{B}}+\mathrm{n}_{+\mathrm{A}} \mathrm{n}_{+\mathrm{B}}\right)
$$

where $F_{-}$and $F_{+}$are estimated from $f(y)$ 's fitted to the sightings to port and starboard respectively. Likewise, $\mathrm{f}_{\mathrm{A}}(0)$ should be replaced by:

$$
\mathrm{f}_{\mathrm{A}}(0)=\left(\mathrm{n}_{-\mathrm{A}} \mathrm{f}_{+\mathrm{A}}(0)+\mathrm{n}_{+\mathrm{A}} \mathrm{f}_{+\mathrm{A}}(0)\right) /\left(\mathrm{n}_{-\mathrm{A}}+\mathrm{n}_{+\mathrm{A}}\right)
$$

where the subscripts - and + refer to data from the port and starboard sides respectively.
(c) Case where $g_{A}(0)$ is desired, $f_{B}(y)$ is not very well known In the IDCR cruises, set $A$ represented the primary observer set used on the entire cruises, while set B was an additional, inferior set used only during the independent observer experiments. It was expected in advance that $g_{B}(0)<g_{A}(0)$, and the value of $g_{B}(0)$ was not of particular
interest. On one of the vessels used, observer $B$ had an obstructed forward view, leading to a rather complex distribution of perpendicular distances with a trough at $\mathrm{y}=0$.

In these cases, estimation of $f_{B}(y)$ is not necessary and may be detrimental due to the extra variance component it introduces. A simpler method of estimating $\mathrm{g}_{\mathrm{A}}(0)$ is to condition on observed perpendicular distribution of B's sightings, i.e.:

$$
\mathrm{F}=\left(1 / \mathrm{n}_{\mathrm{B}}\right) \sum_{\mathrm{i}=1}^{\mathrm{nB}} \mathrm{f}_{\mathrm{A}}\left(\mathrm{y}_{\mathrm{iB}}\right)
$$

where $y_{i B}$ is the perpendicular distance of the ith sighting by observer B.

This simplifies the estimator for $g_{A}(0) / f_{A}(0)$ to:

$$
\left[\mathrm{g}_{\mathrm{A}}(0) / \mathrm{f}_{\mathrm{A}}(0)\right]_{\mathrm{EST}}=\mathrm{n}_{\mathrm{AB}} / \sum_{\mathrm{i}=1}^{\mathrm{n}_{\mathrm{B}}} \mathrm{f}_{\mathrm{A}}\left(\mathrm{y}_{\mathrm{iB}}\right)
$$

(d) Cases where duplicates are not easily identifiable

One of the major problems with the analysis of the earlier parallel ship experiments in the IDCR cruises was that duplicates could not easily be identified. A large number of sightings were classified as possible or probable duplicates in view of their proximity. This problem is reduced by having both observers on the same ship, but apparently some of the decisions about whether observed sightings are duplicates or not are based on the recorded data rather than actual real-time indications such as simultaneous blows (Ward, pers. comm.). In such cases, it may be more appropriate to base the assessment of duplicates on systematic and reproducible criteria at the time of analysis rather than on the hurried judgements made in the field using largely undocumented criteria.

If two pods seen by different observers are widely separated in space or time, then one can be sure that they are not a duplicate. However, if they are sufficiently close together to be a potential duplicate, this fact alone does not imply that they are actually a duplicate, because they may be two separate pods that happen to be rather close together.

The frequency of distinct pods that happen to be rather close together can be assessed from the frequency of such occurrences within each observer's results. Hence the number of duplicates can be estimated by deducting from the observed number of potential duplicate pairs the expected number of pairs of pods that would happen to be close together. The following paragraphs describe the formal steps in the calculation.
(1) List all pairs of sightings. There are $\left(\mathrm{n}_{\mathrm{A}}+\mathrm{n}_{\mathrm{B}}\right)\left(\mathrm{n}_{\mathrm{A}}+\mathrm{n}_{\mathrm{B}}-1\right) / 2$ in all. Divide these into:
(i) pairs of sightings by the same observer ('same pairs') there are $\mathrm{n}_{\mathrm{A}}\left(\mathrm{n}_{\mathrm{A}}-1\right) / 2+\mathrm{n}_{\mathrm{B}}\left(\mathrm{n}_{\mathrm{B}}-1\right) / 2$ of these;
(ii) pairs of sightings by opposite observers ('alternate pairs') - there are ( $\mathrm{n}_{\mathrm{A}} \mathrm{n}_{\mathrm{B}}$ ) of these.
(2) Further subdivide each of the above sets of pairs into 'near' and 'far' according to some threshold of spatial/temporal distance. 'Alternate' pairs in the 'far' category are considered too far apart for there to be any possibility that they are duplicates. The same criterion is used to subdivide the 'same' pairs, although it is assumed that none of the 'same' pairs are actually duplicates. Denote the numbers of pairs in the four categories by NAP, FAP, NSP, and FSP respectively (where F, N stand for 'far' and 'near', and A, S stand for 'alternate' and 'same').
(3) If there were no duplicate sightings, it would be reasonable to assume that the ratio of 'far' to 'near' pairs would be the same amongst the 'same' pairs as amongst the 'alternate' pairs. Because of the presence of duplicates, there will tend to be a higher ratio of 'near' to 'far' pairs among the 'alternate' pairs than among the 'same' pairs. The difference in this ratio gives an estimate of the number of duplicates:

## $\mathrm{n}_{\mathrm{AB}}=$ NAP-FAP.NSP/FSP

Analyses by Cooke (1987a) show that one would not expect the assumption that the ratios of 'far' to 'near' pairs would be the same for 'same' and 'alternate' pairs to hold exactly, but that deviations from the assumption are liable to be slight even in extreme cases.

The variance of this estimator is probably most easily obtained by some non-parametric method such as jackknifing or bootstrapping. Since it is primarily the FSP which provide the information content of the data, such procedures should be conducted conditionally on the observed FSP.

## (e) Case where some sightings are closed on

Independence between observers cannot be maintained if some sightings are closed on. If closure is desirable, then there are three ways to perform independent observer experiments. These are listed below.
(1) Forego closure on part of the cruise (such as alternate transects) and conduct observer experiments only on this part.
(2) Delay all closures until the vessel is abeam of the sighting, and ignore any post-abeam sightings from the analysis; in this case, all the above results apply, although there may be problems of interpretation in high density areas where secondary sightings are occurring.
(3) Close on some of the sightings according to the procedure developed in this section.
Suppose first that closure is performed only on A's sightings. Clearly, once closure begins, B becomes aware of A's sighting. The sightings are then of three kinds:
$\mathrm{n}_{\mathrm{A}^{*}}$ seen by A first, then closed with;
$\mathrm{n}_{\mathrm{B}>\mathrm{A}}$ seen by B first, subsequently by $A$, then closed with; $n_{\mathrm{OB}}$ seen by $B$ only (not closed with).
On the assumption that the two sets of observers have identical detection functions and are independent, then for any given sighting the chances are equal that $A$ or $B$ will see it first. Hence the expected proportion of $\mathrm{n}_{\mathrm{B}>\mathrm{A}}$ sightings is exactly half the expected proportion of duplicates that would occur in the absence of closure. The estimate of $\mathrm{g}_{\mathrm{A}}(0)$ is then

$$
2 \mathrm{f}_{\mathrm{A}}(0) \mathrm{n}_{\mathrm{B}>\mathrm{A}} /\left[\mathrm{F}\left(\mathrm{n}_{\mathrm{OB}}+2 \mathrm{n}_{\mathrm{B}>\mathrm{A}}\right)\right]
$$

All the various other alternative estimates given above for $\mathrm{g}_{\mathrm{A}}(0)$ also apply, with $\mathrm{n}_{\mathrm{AB}}$ replaced by $2 \mathrm{n}_{\mathrm{B}>\mathrm{A}}$.

However, the assumption that each observer is equally likely to be the first to see a sighting is not reasonable for observers occupying different positions on the ship. Violation of this assumption could seriously bias the estimator. To take an extreme example, suppose that observer A searches only the water more than 1 n.miles ahead of the ship, while observer B searches only the water behind this distance. Then $n_{B>A}$ would always be zero.

The only way to ensure equality of detection functions between the closing and non-closing observers is to alternate closure between the two observers. Switching of the closure and non-closure observers should be done after each sighting, whether or not it is closed on. Under these conditions, the estimator for $\mathrm{g}_{\mathrm{A}}(0)$ is

$$
\mathrm{f}_{\mathrm{A}}(0)\left(\mathrm{n}_{\mathrm{A}>\mathrm{B}}+\mathrm{n}_{\mathrm{B}<\mathrm{A}}\right) /\left[\mathrm{F}\left(\mathrm{n}_{\mathrm{OB}}+\mathrm{n}_{\mathrm{A}>\mathrm{B}}+\mathrm{n}_{\mathrm{B}>\mathrm{A}}\right)\right]
$$

where $n_{A>B}$ is the number of sightings seen first by $A$ then by B while closing on B's sightings. An exactly equivalent formula applies for the estimate of $g_{B}(0)$. Note that in the evaluation of this formula, $f_{A}(y)$ is the pdf of perpendicular sightings distances for those of A's sightings obtained while closing on $A$, and $f_{B}(y)$ is the pdf of perpendicular sightings distances of those of B's sightings obtained while closing on B.
(f) Case where the independence assumption does not hold If the assumption of independence between the sightings of a pod by the two observers is violated, this could lead to a bias in the $g(0)$ estimates if there is a net positive or negative correlation between the sighting probabilities. One way to reduce the bias would be to stratify the data by factors thought to influence the sighting probabilities (other than perpendicular distance, which has already been taken into account) such as sighting conditions. This will reduce the bias at the expense of increasing the variance of the $g(0)$ estimate. Since population estimates are proportional to $1 / \mathrm{g}(0)$, these are liable to small sample bias even when the estimates of $g(0)$ are unbiased. With sufficiently fine stratification, this small sample bias could become very large: if some strata contain no duplicate sightings, the population estimate will be undefined. Small sample bias could be reduced (and reversed in direction) by replacing ratios such as $n_{B} / n_{A B}$ by $\left(1+n_{A}\right) /\left(1+n_{A B}\right)$.

Alternatively, a positively biased estimate of $g_{A}(0)$ can be obtained on the assumption that $\mathrm{g}_{\mathrm{A}+\mathrm{B}}(0)=1$.

### 4.3 Whales or pods as targets?

In line transect sampling for cetaceans there is usually no alternative to the use of pods as targets because it is impossible to estimate the number of different whales sighted from the platform as it moves along the transect. This raises a number of related problems: pod definition (e.g. Cooke, 1986b), estimating the size of detected pods (e.g. Butterworth and McQuaid, 1986), relating the size of detected pods to mean pod size in the population, dependence of the shape of $g(y)$ on pod size. The question of pod definition is complex in those cases where pod structure is fluid (Gordon, 1987; Whitehead, 1985) or difficult to discern on the basis of behaviour or spatial association. It is not considered further here except to note that if the sampling procedure used to estimate $g(0) / f(0)$ for pods is distinct from that used to estimate population mean pod size, as is the case in recent IDCR cruises, then there is a risk of different pod definitions being used in the different procedures.

In general, detection probabilities will be related to pod size, with sighting curve $\mathrm{g}_{s}(\mathrm{y})$ applying to pods of size s. For the pooled population

$$
g(y)=\sum_{s} P_{s} g_{s}(y) \text { and esw }=\sum_{s} P_{s} e_{s} w_{s}
$$

where $P_{s}$ is the true (unknown) proportion of pod size $s$ in the population (Quinn, 1979). If the size of each pod detected is estimated, then one way to estimate total abundance of pods is to post-stratify on pod size and sum
the estimates for each size; the sum is unbiased if the estimate for each pod size is unbiased. If size estimates for each pod are not available the only alternative is to fit to the pooled data - the two methods are, by definition, equivalent if the model used for $g(y)$ is 'pooling robust' (Burnham et al., 1980, p.45). Quinn (1985) shows that the CV of the estimated number of pods by pooling is always less than or equal to that for the method of post-stratification. However, whereas from the post-stratification method the estimate of whale abundance is simply the sum of pod abundance multiplied by pod size over all sizes, for the pooling method an estimate of mean pod size is required from the data. Given the size of each detected pod and a pooling robust estimator, an unbiased estimate of mean pod size in the population is the mean of the observed pod size weighted by $1 / \mathrm{esw}_{\mathrm{s}}$ (Quinn, 1979). An alternative is to use the mean size of pods detected 'close' to the trackline, or to use a linear regression of observed pod size against $y$ evaluated at $y=0$ (e.g. Best and Butterworth, 1980; Hammond, 1986b); however, those methods may be biased as a result of the dependence of $\mathrm{g}(0)$ on pod size (Cooke, 1985).

Another alternative is to take whales, rather than pods, as the primary targets. That is, interpret the detection of a pod as the detection of each of its members, at the same $y$ value as estimated for the pod. In practice this means using the sum of the estimated sizes of all detected pods for the total number of targets detected, $n$, and applying the chosen estimator for $f(0)$ to the $y$ values for each pod weighted by pod size. Programs available for estimation of esw do not permit weighting of data points in this way; however, the same result is obtained by repeating the data value which would normally apply once to a pod of size $S, S$ times. If the estimator used for $f(0)$ is fully pooling robust, the result of this procedure is identical to multiplying the estimate of $f(0)$ from the pooled, unweighted data on pod sightings by the estimate of mean pod size derived by Quinn, i.e. the mean of the observed pod sizes weighted by 1/esws.

The whales-as-targets approach does not appear to have been recommended in the literature, presumably because violation of the assumption that detection probabilities are independent means that the usual estimate of the variance of the estimator chosen for $f(0)$ is no longer appropriate. Estimation of the variance of $f(0)$ would have to be based on the jackknife or bootstrap approach.

Another problem in relation to cetacean surveys is that $g(0)$ estimates from independent observer experiments are derived from the proportion of pods detected as duplicates, and thus apply to pods, not whales. However, $g(0)$ for whales should be larger than for pods because pods missed on the trackline will tend to be small, so that the proportion of whales missed will be less than the proportion of pods missed. One solution, therefore, would be to use the value of $g(0)$ estimated for pods pending development of an estimator for whales.

There may also be concern about the size of the variance of $f(0)$ from the whales-as-targets method. However, the variance of the estimate of whale abundance may be no higher than that based on fitting $f(y)$ to pods. To take the simplest example, suppose whale density was to be estimated by fitting a negative exponential model for $f(y)$ to the distribution of distance to sighted pods, and multiplying the usual estimator of pod density by the mean observed pod size

$$
(n / 2 L \bar{y}) \cdot(1 / n) \cdot \sum_{i=1}^{n} S_{i}=\sum_{i=1}^{n} S_{i} / 2 L \bar{y}
$$

Taking each whale as a primary sighting and using the same estimator for the mean density of those targets gives

## $\Sigma \mathrm{S}_{\mathrm{i}} /\left(2 \mathrm{~L} \Sigma \mathrm{y}_{\mathrm{i}} \mathrm{S}_{\mathrm{i}} / \Sigma \mathrm{S}_{\mathrm{i}}\right)$

because the mean $y$ for whales is the mean $y$ for pods weighted by pod size. The second estimator of whale density has lower variance than the first because, assuming the larger $y$ values tend to be for the larger pods, the total number of whales detected, $\Sigma \mathrm{S}_{\mathrm{i}}$, will be more strongly correlated with the weighted mean $\Sigma y_{i} \mathrm{~S}_{\mathrm{i}} / \Sigma \mathrm{S}_{\mathrm{i}}$ than with the unweighted mean $\bar{y}$.

We would recommend that, when possible, this method of fitting $f(y)$ to the distribution of $y$ values for pods, weighted by pod size, be investigated. It may be difficult to obtain pod size estimates when following the chosen procedure for making primary sightings, as is the case for passing mode transects. However, the following points are relevant: (1) as for distance estimation and species identification, errors in pod size estimation for pods far from the trackline have little effect on abundance estimation; (2) pod size estimates derived from different sampling procedures may be inappropriate because of differences introduced by problems in pod definition; and (3) it may be possible to apply a correction for consistent errors in pod size estimation.

Recent approaches by Drummer and McDonald (1987) and Ramsey, Wildman and Engbring (1987) are relevant to the consideration of the effect of varying pod size on abundance estimation.

## 5. MODELLING PERPENDICULAR DISTANCE DATA FROM LINE TRANSECT SURVEYS

### 5.1 Definition of the problem

Perpendicular distance data calculated from sighting angles and distances collected during IWC/IDCR Antarctic cruises are suitable for investigating the fit of different models because they typically contain many values at or close to zero; a problem common to many data

Table 1
Perpendicular distance frequencies for the data set of Figs 2-7

| Perpendicular <br> distance <br> interval (n.miles) | Fig. 2 | Fig. 4 | Fig. 5 | Fig. 7 |
| :--- | :---: | :---: | :---: | :---: |
|  | $27 *$ | 22 | $69 * *$ | 63 |
| 0.1 | 0.2 | $11 *$ | 17 | $45 * *$ |
| $0.2-0.3$ | 16 | 16 | $39 * *$ | 49 |
| $0.3-0.4$ | 17 | 16 | 27 | 39 |
| 0.4 | 0.5 | 16 | 17 | 20 |
| $0.5-0.6$ | 17 | 14 | 23 | 24 |
| 0.6 | 0.7 | 14 | 11 | 19 |
| 0.7 | 0.8 | 6 | 9 | 15 |
| 0.8 | 0.9 | 9 | 7 | 14 |
| $0.9-1.0$ | 4 | 6 | 16 | 18 |
| 1.0 | 1.1 | 4 | 5 | 19 |
| $1.1-1.2$ | 12 | 5 | 15 | 16 |
| $1.2-1.3$ | 3 | 5 | 8 | 15 |
| $1.3-1.4$ | 3 | 4 | 2 | 14 |
| $1.4-1.5$ | 4 | 4 | 15 | 11 |
| $1.5-1.6$ | 3 | 3 | 7 | 8 |
| $1.6-1.7$ | 1 | 3 | 5 | 8 |
| $1.7-1.8$ | 2 | 2 | 9 | 7 |
| $1.8-1.9$ | 0 | 2 | 4 | 6 |
| $1.9-2.0$ | 1 | 8 | 5 |  |

[^6]

Fig. 2. perpendicular distance data, grouped by 0.1 nm intervals, from the southern stratum of Area IV W, 1984/5. Vessel was SM2.
(a) Fit of negative exponential model.
(b) Fit of exponential power series (or generalised exponential) model.
(c) Fit of hazard-rate model.
(d) One-term fit of Fourier series model. This is the fit selected by sequential likelihood ratio tests. (e) Two-term fit of Fourier series model.
(f) Three-term fit of Fourier series model. This is the optimum overall fit of the Fourier series model, as assessed by the likelihood ratio tests.
(g) Four-term fit of Fourier series model.
(h) One-term fit of Hermite polynomial model. This is the optimum fit of the Hermite polynomial model, as assessed by the likelihood ratio tests.
(i) Two-term fit of Hermite polynomial model.
(j) Three-term fit of Hermite polynomial model.
(k) Four-term fit of Hermite polynomial model.
sets. Such data are difficult to model. The shape criterion of Burnham et al. (1980, p.47) suggests that a line transect model should have a 'shoulder', in the sense that a pod of whales at perpendicular distance 0.2 n .miles say from the vessel track line should be almost as likely to be detected as a pod located on the track line. (The curve of Fig. 2(c) provides a good illustration of a shoulder.) The data frequently appear to violate this requirement and, in the absence of satisfactory models, estimates of minke whale stocks in the Antarctic have generally been generated from simplistic and unrealistic models of the perpendicular distance distribution. For example, the estimates tabulated in IWC (1986b) were calculated by fitting a negative exponential distribution to the data. Although in principle we may multiply any resulting estimate by a correction factor, $e$, to compensate for the poor choice of model, in
practice we are unable to estimate e reliably; furthermore, e is unlikely to be constant.

Here, we examine the problem of modelling the perpendicular distance data through detailed analyses of two data sets. The first was recorded from vessel SM2 in the southern stratum of Area IVW during 1984/85. Angle boards were used, and sighting angles and distance appear to be relatively accurate for this data set (Buckland and Anganuzzi, 1988). The data are relatively straightforward to interpret, and we are therefore able to contrast model performance in a meaningful way. The second data set was recorded from vessel T11 in Area V during 1980/1, and is more typical of the data sets up to and including the 1983/4 season. The perpendicular distance frequencies for both sets are given in Table 1. Those of the first set are featured in Figs 2-4, and those of the second in Figs 5-7.


Fig. 3. Perpendicular distance data of Fig. 2, with first two groups combined. (d) This is the fit selected by sequential likelihood ratio tests. (e) This is the optimum overall fit of the Fourier series model, as assessed by likelihood ratio tests. (f) This is the optimum fit of the Hermite polynomial model, as assessed by likelihood ratio tests.



Fig. 4. Perpendicular distance data of Fig. 5.2, smeared before analysis. (d) This is the optimum fit of the Fourier series model, as assessed by likelihood ratio tests. (e) This is the optimum fit of the Hermite polynomial model, as assessed by likelihood ratio tests.

### 5.2 Investigation of a 'good' data set

The histogram of Fig. 2 suggests that these data derive from a detection curve possessing a shoulder that extends out to roughly 0.6 n.miles. The relative abundance of records in the first interval (27) and the corresponding sparsity in the second interval (11) probably reflect rounding errors in the data and in particular, rounding of small angles to zero which leads to a calculated value of zero for the corresponding perpendicular distances. The average frequency for these two intervals (19) is close to those recorded in the next four intervals $(16,17,16,17$ respectively). In Fig. 3, the same data are illustrated, but
with the first two intervals combined, in an attempt to reduce the effects on the line transect models of rounding errors in the data. Smearing, as described by Butterworth (1982a), has traditionally been used on minke whale sightings data for the same purpose and in Fig. 4 we show the same data set after smearing. The smearing method of Hammond and Laake (1983) was adopted, but with a smearing sector defined by $\theta \pm 7.5^{\circ}$, and $\mathrm{r} \pm 0.25 \mathrm{n}$. miles, where $\theta$ is the sighting angle and $r$ the sighting distance. (Buckland and Anganuzzi (1988) show that choice of smearing technique, within reason, has little effect on the analysis.) The smearing method eliminates completely the


Fig. 5. Perpendicular distance data, grouped by 0.1 nm intervals, recorded by vessel T11, 1980/1. (d) This is the optimum fit of the Fourier series model, as assessed by likelihood ratio tests. (e) This is the fit selected by sequential likelihood ratio tests. (f) This is the optimum overall fit of the Hermite polynomial model, as assessed by likelihood ratio tests.
'heap' of records, caused by rounding error, between 1.2 n .miles and 1.3 n .miles perpendicular distance. It also retains the shoulder that these data exhibit, although the number of observations in the first interval remains rather high relative to the number in subsequent intervals.

### 5.2.1 Comparison of models

The models fitted to these three representations of the data are those investigated by Buckland (1987): the negative exponential; the exponential power series, or generalised exponential; the hazard-rate; the Fourier series; and the Hermite polynomial. The negative exponential model provides almost identical estimates of $f(0)$, from which the estimated density of whales is calculated, for each of the three representations, but overestimates relative to the other models. Further, its ability to fit the data adequately is dubious. The exponential power series fits exhibit a small shoulder, and rather variable estimates of $f(0)$. The hazard-rate fits match the above interpretation of the data reasonably well, with the flat section of the shoulder extending out to around 0.3 n .miles in all three cases. To apply either the Fourier series or the Hermite polynomial, it is necessary to decide on the number of terms to use. We adopt here likelihood ratio tests. However, we may utilise them in more than one way. We may test sequentially, starting at the one-term model, and stop when we fail to obtain a significant improvement in the fit. Alternatively, we may fit the model with one, two, three ... terms, up to some upper limit, taken to be four terms here, and use likelihood ratio tests to select the optimum fit. We use both approaches; where they indicate different numbers of terms, we plot both. In Fig. 2 only, all fits up to the four-term model are shown for both the Fourier series and the Hermite polynomial models.

Buckland (1985) found that the Fourier series model performs badly on data sets with a small or no shoulder, with the estimate of $f(0)$ tending to increase as the number of terms fitted increases. In Figs 2(d) to 2(g), we see evidence of this effect, although the data appear to exhibit a wide shoulder. If we use sequential likelihood ratio testing, our estimate of $f(0)$ is 0.95 . Taking the optimal fit, as described above, we have the estimate $f(0)=1.24$. The latter estimate leads to an estimated number of whales more than $30 \%$ higher than the former. Such inconsistency for data that are relatively very well behaved seems unacceptable. In Fig. 3, the discrepancy is about 14\%, and if we smear the data before analysis, both rules select the two-term fit, and we obtain an estimate almost identical to that of the hazard-rate model. Hence, if we reduce the effect of rounding errors on these data, the Fourier series model performs better. The behaviour of the Hermite polynomial model is similar to that of the Fourier series, except that the number of terms used is less highly correlated with the estimate of $f(0)$ (Buckland, 1985 and Figs 2(h) to 2(k)). Further, for a given number of terms, it tends to provide a better fit to data than the Fourier series model and, when the number of terms is small, it yields a more reliable (and larger) estimate of variance than the Fourier series (Buckland, 1985; 1987). Hence, for this data set, we might have reservations about using the Fourier series model. In addition, we would expect to overestimate whale density if we use the exponential model, unless we can correct for bias. The other three models probably perform adequately on these data.

### 5.3 Investigation of a poor data set

Unfortunately, data from earlier years are more problematic. For example, the data recorded from vessel T11 in 1980/1 show no existence of a shoulder before or after smearing (Figs 5-7).


Fig. 6. Perpendicular distance data of Fig. 5, with first three groups combined.
(a) Fit of negative exponential model.
(b) Fit of exponential power series model.
(c) Fit of hazard-rate model.
(d) Four-term fit of Fourier series model. This is the optimum fit of the Fourier series model, as assessed by likelihood ratio tests.
(e) Two-term fit of Hermite polynomial model. This is the fit selected by sequential likelihood ratio tests.
(f) Four-term fit of Hermite polynomial model. This is the optimum overall fit of the Hermite polynomial model, as assessed by likelihood ratio tests.


Fig. 7. Perpendicular distance data of Fig. 5, smeared before analysis.
(a) Fit of negative exponential model.
(b) Fit of exponential power series model.
(c) Fit of hazard-rate model.
(d) Four-term fit of Fourier series model. This is the optimum fit of the Fourier series model, as assessed by likelihood ratio tests.
(e) Two-term fit of Hermite polynomial model. This is the fit selected by sequential likelihood ratio tests.
(f) Four-term fit of Hermite polynomial model. This is the optimum overall fit of the Hermite polynomial model, as assessed by likelihood ratio tests.

Most of the perpendicular distances recorded as zero correspond to sighting distances of 2.5 n .miles or less. Suppose we assume that the true perpendicular distances were at most 0.3 n . miles, and hence group the first three intervals (Fig. 6). Angles greater than around $7^{\circ}$ would have to be rounded to zero to violate this assumption. Either accuracy in measuring angles is very poor or the data have a very narrow shoulder.

### 5.3.1 Comparison of models

The negative exponential model is the only model that yields smaller estimates of $f(0)$, and hence larger estimates of effective track width, for these data than for the first data set. In view of the absence of a shoulder for these data, and since the first data set is from an ice-edge stratum in which visibility was relatively good, it seems highly improbable that the estimates from the negative exponential model are consistent in their bias. This implies that, if the negative exponential model is adopted, not only should there be a correction for bias, but also the correction should be different for different data sets. In view of the difficulties in estimating this correction even when it is assumed constant, the use of the negative exponential model is of doubtful value. However, there is not a clear alternative. Figs 5(b), 6(b) and 7(b) show that estimates under the exponential power series model are highly unstable. For data such as these, the value of the likelihood function under the exponential power series model is insensitive to variation in $f(0)$, so that we obtain very imprecise estimation. In addition, the model may overestimate dramatically the number of whales in an area. Hence, although the model generally provides good fits to minke data (Buckland, 1987) it generates estimates that are both very imprecise and highly biased. Of the other four models considered here, only the hazard-rate is able to fit data sets of the kind illustrated in Figs 5-7 as well as the exponential power series. It always has a shoulder, and so avoids the major problem of the exponential power series model. However, the advantage of the hazard-rate model can also be its disadvantage. If the spike exhibited by most of the minke data sets is an artifact of rounding error or bias, then a model with the flexibility to fit the spike will not provide robust analyses of uncorrected data. If we can first successfully 'correct' the data, for example by smearing, then the hazard-rate model should be a powerful and reliable tool for providing estimates of numbers of minke whales. However, as yet, there has been no convincing explanation of the disproportionately large spike in many of the data sets, and smearing has been only partially successful.

The hazard-rate model yields fitted curves that are smooth and non-increasing. Four-term fits of the Fourier series and Hermite polynomial models show undulations, where they attempt to follow humps and hollows in the data, yet still fail to provide such good fits to the data, as assessed by the likelihood or by the $\chi^{2}$ goodness-of-fit statistic. The estimates for both models are little affected by smearing (Buckland, 1987, and Figs 5-7), but are highly dependent on the number of terms fitted (compare for example Fig. 7(e) with Fig. 7(f)).

### 5.4 Discussion

Most analyses of the Southern Hemisphere minke sightings data have assumed the negative exponential model. It is
usually acknowledged that resulting estimates are likely to be biased, but little attention has been paid to the likelihood that the bias is variable. The first data set considered here was collected from the ice-edge stratum, during usually good sighting conditions, and exhibits a wide shoulder. In a northern stratum, frequently.with poor sighting conditions, we might expect the shoulder to be substantially more narrow. The negative exponential model has a single, inflexible shape, and therefore leads to different biases in the two cases. To avoid this problem, we require an estimator that is model robust (Burnham et al., 1980, p.44); we require a model that has the flexibility to fit the different shapes that the true detection function exhibits for different data sets. For the minke data, the Fourier series estimator, in conjunction with likelihood ratio testing to select the number of terms, is not model robust, and if we allow more terms to enter, the fitted detection curve exhibits implausible shapes. Similarly, the Hermite polynomial estimator is not particularly model robust in this context. Hayes and Buckland (1983) state: 'In general, an estimator that is not model robust will provide an estimate whose standard error is misleadingly small'. We can therefore expect biased estimation and standard errors that are too small from the negative exponential and both series-type models. The hazard-rate and the exponential power series estimators are both model robust and, for precisely this reason, are sensitive to rounding errors in the data, which change the apparent shape of the detection curve. The exponential power series model may be ruled out since it gives very unstable estimates, often with high positive bias, when data show little or no evidence of a shoulder. This leaves us with the hazard-rate model, which is likely to be satisfactory provided we have a smearing technique that works.
For many minke data sets, smearing fails to eliminate the spike in the perpendicular distances at zero distance. It is useful, therefore, to consider how this spike might arise. Some possible explanations are listed below.
(a) The spike may be real. Evidence suggests that the probability of detection of a pod of minke whales on the trackline, $g(0)$, is well below unity. However, it may be that minkes, whose behaviour makes them difficult to detect at 100 metres or so from the vessel, are more easily detectable at 10 or 20 metres. In other words, $g(0)$ may equal, or be very close to, unity, but the probability may fall away very quickly as distance from the trackline increases. If this is the case, we must explain why the 1984/5 data sets fail to support this hypothesis.
(b) There may be a tendency to underestimate small sighting angles. The absence of a spike in the 1984/5 data sets, when angle boards were used, perhaps favours this explanation. Although smearing appears to correct rounding of angles to zero reasonably successfully (Buckland and Anganuzzi, 1988), it is not designed to correct for consistent underestimation of angles, in which for example an angle of $9^{\circ}$ may tend to be recorded as $5^{\circ}$.
(c) The first sighting of many pods may be of very brief duration. If the observer first scans, to attempt to resight the pod to verify his sighting, but fails to locate it immediately, the original sighting angle he records may be very imprecise. The coarser the scale to which angles are rounded, the more spiked the data will be.
(d) A proportion of whales may be attracted to the vessel, hence creating the observed spike. This effect is known to occur in surveys of some populations of dolphins and porpoises.
(e) The whales are assumed to be stationary. In practice, they are moving, and many of the pods will move across the bow of the vessel. It is conceivable that a whale, sensing the approach of a vessel, increases speed. This in turn may lead to an increase in the number of blows, or an increase in the disruption of the sea surface. In other words, whales on or near the vessel trackline may be more detectable at a given sighting distance than whales at larger sighting angles. When combined with random whale movement, this will lead to an artificial spike in the data, and to overestimation of whale densities.

It is clear that further investigation of the above points is necessary before stronger recommendations on how to model the data can be made.

## 6. EFFECTS OF WHALE MOVEMENT

The implications of various types of whale movement for the estimation of abundance or density have received relatively little consideration in the literature, consequently much of the material in this chapter is new. The increase in 'encounter' rate due to random target movement was considered by Yapp (1956) and Skellam (1958). Further results relevant to cetacean survey include Hiby (1982), Basson and Butterworth (1984), Cooke (1985) and Kishino (1986). The detection and implications of whale movements in response to the platform are considered in Smith (1979) and Burnham et al. (1980). Schweder (1977) developed a general formalism which encompasses both random and directed target motion. The methods used below are more suited to our limited mathematical background.

This chapter considers the effects of whale movements under three main categories: effects on estimates derived when successive transects are located independently; the effects of whale movement on estimates derived from survey 'grids'; and the detection and effects of whale response to platforms.

### 6.1 Transects located independently

For consideration of whale movement the definitions of $\mathrm{p}^{\prime}{ }_{i}$ and esw in terms of stationary whales (Section 2.1) are inadequate and we define new parameters $\mathrm{p}_{\mathrm{i}, \mathrm{m}}$ and $\mathrm{esw}_{\mathrm{m}}$ in terms of the position at which the target crosses the abeam line: Let $\mathrm{G}_{\mathrm{i}, \mathrm{m}}(\mathrm{y})$ be the pdf for the distance, y , at which the ith target crosses the abeam line. Define $g_{m}(y)$ as the probability of detection for a target which is on course to cross the abeam line at $y$. Thus the unconditional probability of detection for the ith target

$$
p_{i, m}=\int_{0}^{\infty} G_{i, m}(y) g_{m}(y) d y
$$

Assuming all detected targets are on course to cross the abeam line within distance $w$, and $G_{i, m}(y)$ is constant at $\mathrm{p}^{\prime}{ }_{i, m}$ for $0<\mathrm{y}<\mathrm{w}$,

$$
p_{i, m}=p_{i, m}^{\prime} \int_{0}^{w} g_{m}(y) d y
$$

We define $\operatorname{esw}_{\mathrm{m}}$ as: $\int_{0}^{w} g_{m}(y) d y$
Target motion then affects the estimate of abundance derived from an analysis based on stationary targets to the extent that the product $\mathrm{p}^{\prime}{ }_{\mathrm{i}, \mathrm{m}}$. esw $_{\mathrm{m}}$ differs from the product $\mathrm{p}_{\mathrm{i}}^{\prime}$.esw estimated for stationary targets. Estimates of local density may be affected by the difference between esw ${ }_{m}$ and the estimate of esw, and by the distortion in the measure of effort, $L$, which occurs if whale movement is generally either in the same or the opposite direction as the movement of the survey vessel (see below).

Consider, first, the effect of random whale movement; that is, there is a uniform distribution in the direction of travel. There will be little or no difference, on average, between the $\mathrm{p}_{\mathrm{i}}^{\prime}$ values calculated and $\mathrm{p}_{\mathrm{i}, \mathrm{m}}^{\prime}$, nor will the measure of effort, L , be distorted. Effects on estimates of density and abundance are therefore through esw, and are the same for both. There are two effects. Firstly, esw $_{m}$ exceeds esw to an extent depending on the speed of whale movement relative to vessel speed, the value of $g(0)$ and the degree to which effort is allocated forward, along the trackline. This is the effect referred to in Hiby (1982) as the 'increase in sighting rate'. Secondly, the shape of the distribution of $y$ distances to sightings is distorted resulting in a downward bias in the estimate of esw, unless strip transect sampling is used or the function chosen to fit $f(y)$ is insensitive to shape. These effects combine to give an upward bias to estimates of abundance or density. The size of the bias is not likely to be significant for most platform/whale speed combinations; however, the simulation study by Hiby (1982) indicated an upward bias of about $50 \%$ in the case of equal platform and whale movement speeds which argues strongly against the use of LT sampling from very slow platforms. It also means that estimates based on 'Kelker strip' analysis of LT data are not necessarily downward biased.
Consider now the effects of whale movement which is not random but is generally in one direction. The effect of local movements perpendicular to the trackline would be expected, on arguments based on an encounter region model, to be about $50 \%$ larger than that due to 'random' movement. Movement parallel to the trackline has no effect on esw. Density estimates are affected via distortion of the effort measure L. Estimates are biased down by a factor ( $\mathrm{v}-\mathrm{w}$ )/v for whale movement in the same direction as platform movement and up by $(v+w) / v$ for whale movement in the opposite direction (assuming $g(0)$ close to $1, \mathrm{v}=$ platform speed, $\mathrm{w}=$ whale speed). There is no effect on abundance estimates if there is no movement across the boundaries of the area surveyed because $\mathrm{p}_{\mathrm{i}}$ values for whales within the area are unaffected (in other words, we observe a higher density over a shorter distance or a lower density over a longer distance).

Estimates of esa using the cue counting approach (Section 2.5) are unaffected by any type of target movement, so that estimates of local density by this method are also unaffected. Conversely, estimates of abundance (see Section 8) will be biased by target movement parallel to the trackline.

### 6.2 Effects of whale drift across a grid survey

Serious bias in estimates of abundance is more likely to occur via interaction between whale movement and the spacing of successive transects in a grid or saw-tooth
pattern for shipboard survey. The problem arises if there is a general drift in the whale population which either opposes or complies with the direction of progress of the survey. This results in the platform spending either too little or too much time in the midst of the whales. In terms of our model for estimating abundance, the calculations of $\mathrm{p}_{\mathrm{i}}^{\prime}$ fail to reflect the average values of $\mathrm{p}_{\mathrm{i}, \mathrm{m}}$ in the population. To quantify the error consider the possible values of $\mathbf{p}^{\prime}{ }_{i, m}$ for whales which are moving along an iso $\mathrm{p}_{\mathrm{i}}$ line.


Fig. 8. Diagram of whale drift problem. Vessel progresses west to east at Wnm per day, whale drifts east to west at dnm per day.

For example, in a given stratum the survey design used may assign equal $\mathrm{p}^{\prime}{ }_{\mathrm{i}}$ values to all points along a certain latitude. Suppose the survey has a random start at the western boundary of the stratum and follows a pattern of roughly north-south transects which cross the latitude once a day, each crossing W miles further to the east than the previous one. Let x represent the distance along the line of latitude from the western boundary (see Fig. 8). Let the first transect cross the line at a randomly chosen point $\mathrm{x}_{\mathrm{v}, 0}$ between O and W . Thus on the kth day the kth transect crosses the line at $\mathrm{x}_{\mathrm{v}, 0}+\mathrm{kW}$. Suppose whale i is at $\mathrm{x}_{\mathrm{w}, 0}$ on day 0 and moves westwards at $d$ miles per day so that at the time the kth transect crosses the line the whale is at $\mathrm{x}_{\mathrm{w}, 0}-\mathrm{kd}$. Thus the whale will come within $l \mathrm{n}$. miles of the transect if, for some integer $k$,

$$
\left(\mathrm{x}_{\mathrm{v}, 0}+\mathrm{kW}\right)-\left(\mathrm{x}_{\mathrm{w}, 0}-\mathrm{kd}\right)<l
$$

where $l$ equals $1 / \sin \theta$ and $\theta$ is the angle at which the transect crosses the line of latitude. This conditional is equivalent to

$$
\left|\mathrm{x}_{\mathrm{v}, 0^{-}}\left(\mathrm{x}_{\mathrm{w}, 0^{-}} \mathrm{k}(\mathrm{~W}+\mathrm{d})\right)\right|<l
$$

i.e. the randomly chosen position for the first crossing lies within distance $l$ of the point $\mathrm{x}_{\mathrm{w}, 0}-\mathrm{k}(\mathrm{W}+\mathrm{d})$ for some integer k. Consider Fig. 9. With this choice of $\mathrm{x}_{\mathrm{w}, 0}, \mathrm{x}_{\mathrm{v}, 0}$ has probability $2 l / \mathrm{W}$ of lying within $l$ of the point $\mathrm{x}_{\mathrm{w}, 0}-\mathrm{k}(\mathrm{W}+\mathrm{d})$ so that

$$
\mathrm{p}_{\mathrm{i}, \mathrm{~m}}^{\prime}=\mathrm{p}_{\mathrm{i}}^{\prime}=2 l / \mathrm{W} .
$$

Now move the initial whale position $\mathrm{x}_{\mathrm{w}}$ westwards from $\mathrm{x}_{\mathrm{w}, 0}$ a distance $\mathrm{x}_{\mathrm{w}, 0}-\mathrm{k}(\mathrm{W}+\mathrm{d})-l$ until the position in Fig. 9 is reached.

Throughout that westward shift the value of $\mathrm{p}_{\mathrm{i}, \mathrm{m}}^{\prime}$ remains $2 l / \mathrm{W}$ but the probability is reduced to zero if $\mathrm{x}_{\mathrm{w}}$ is shifted westwards for a further distance $2 l$ (Fig. 9). The average value of $\mathrm{p}_{\mathrm{i}, \mathrm{m}}$ for this latest shift is thus $1 / \mathrm{W}$. $\mathrm{p}^{\prime}{ }_{i, m}$ remains at zero as $\mathrm{x}_{\mathrm{w}}$ is shifted westwards by a further distance $\mathrm{W}+\mathrm{d}-2 l$, then increases to $2 l / \mathrm{W}$ during a further shift of $2 l$ to attain the position in Fig 9.

Finally, by shifting a further $\mathrm{W}-\mathrm{l}-\left(\mathrm{x}_{\mathrm{w}, 0}-\mathrm{k}(\mathrm{W}+\mathrm{d})\right.$ ) westwards the original position is regained with

$$
\mathrm{x}_{\mathrm{w}}-(\mathrm{k}-l)(\mathrm{W}+\mathrm{d})=\mathrm{x}_{\mathrm{w}, 0}-\mathrm{k}(\mathrm{~W}+\mathrm{d})
$$

The value of $\mathrm{p}_{\mathrm{i}, \mathrm{m}}$ remains at $2 l / \mathrm{W}$ throughout the final shift, thus the average value of $\mathrm{p}_{\mathrm{i}, \mathrm{m}}$ is

$$
\begin{aligned}
& \frac{\mathrm{x}_{\mathrm{w}, 0}-\mathrm{k}(\mathrm{~W}+\mathrm{d})-l}{\mathrm{~W}+\mathrm{d}} \frac{2 l}{\mathrm{~W}}+\frac{2 l}{\mathrm{~W}+\mathrm{d}} \cdot \frac{l}{\mathrm{~W}}+\frac{\mathrm{W}+\mathrm{D}-2 l}{\mathrm{~W}+\mathrm{d}} \cdot \\
& +\frac{2 l}{\mathrm{~W}+\mathrm{d}} \cdot \frac{l}{\mathrm{~W}}+\frac{\mathrm{W}-l-\left(\mathrm{w}_{\mathrm{x}, 0}-\mathrm{k}(\mathrm{~W}+\mathrm{d})\right)}{\mathrm{W}+\mathrm{d}} \cdot \frac{2 l}{\mathrm{~W}} \\
& =2 l(\mathrm{~W}+\mathrm{d})=\mathrm{p}_{\mathrm{i}}^{\prime} \mathrm{W} /(\mathrm{W}+\mathrm{d}) .
\end{aligned}
$$

A similar argument shows that if the whale moved eastwards (in the same direction as the progress of the survey) then $\mathrm{p}^{\prime}{ }_{\mathrm{i}, \mathrm{m}}$ increases to

$$
2 l /(\mathrm{W}+\mathrm{d})=\mathrm{p}_{\mathrm{i}}^{\prime} \mathrm{W} /(\mathrm{W}-\mathrm{d})
$$

Thus, on average, $\mathrm{p}^{\prime}{ }_{\mathrm{i}}$ overestimates $\mathrm{p}^{\prime}{ }_{\mathrm{i}, \mathrm{m}}$ by a factor ( W $+d) / W$ if whale i moves against the progress of the survey and underestimates $\mathrm{p}^{\prime}{ }_{\mathrm{i}, \mathrm{m}}$ by a factor (W-d)/W if whale i moves with the progress of the survey. Estimates of abundance are thus biased downward by $\mathrm{W} /(\mathrm{W}+\mathrm{d})$ and upward by $\mathrm{W} /(\mathrm{W}-\mathrm{d})$, respectively.


Fig. 9. The interaction of whale and vessel positions, as described in the derivation of Section 6.2. The numbered stars represent examples of the quantity

$$
\begin{gathered}
\mathrm{x}_{\mathrm{w}, 0}-\mathrm{k}(\mathrm{~W}+\mathrm{d}) . * 1, * 2 \text { and } * 3 \text { represent } \mathrm{x}_{\mathrm{w}, 0}-(\mathrm{k}+1)(\mathrm{W}+\mathrm{d}) \\
\mathrm{x}_{\mathrm{w}, 0}-\mathrm{k}(\mathrm{~W}+\mathrm{d}) \text { and } \mathrm{x}_{\mathrm{w}, 0}-(\mathrm{k}-1)(\mathrm{W}+\mathrm{d}) \text { respectively. }
\end{gathered}
$$

The derivation of these bias factors has been presented in detail because the effect can be considerable. This is due to the fact that the speed of whale movement is related to the rate of progress of the survey, not to the speed of the platform. For example, suppose the concentration of whales within the survey area drifts West to East at a half knot, i.e. 12 n.miles per day. Then two surveys, carried out at a rate of 50 n .miles per day, one from East to West, the other from West to East, will differ in expected estimate of abundance by a factor of $(50+12) /(50-12)$, or approximately $5 / 3$.

If it is suspected that a shift in distribution of the population may occur in a certain direction, for example, that the population may move slowly along a ridge, the progress of the survey should be perpendicular to that suspected movement with individual transects parallel to it. Unfortunately most shifts in distribution will not be predictable. These effects could be regarded as a component of variance; however, that component will not be reflected in the estimate of variance obtained from replicate transects - to take an extreme example, if the survey keeps pace with a concentration of whales, abundance will be greatly overestimated with a low variance estimate.

A related question concerns the deployment of a number of platforms to survey an area. On a purely intuitive basis, we would recommend that the survey be planned to avoid adjacent areas being surveyed at different times, which would risk a concentration being counted either twice or not at all, if it moved from one area to the other between surveys. The survey should therefore be planned to have vessels diverging from and converging to one or more 'nodes'.

### 6.3 Movement in reaction to platforms

Movement in reaction to a survey platform will bias the estimated esw through changes in the distribution of perpendicular distances. This only occurs if such movement occurs before detection of the whales of pods. Reactive movement occurring after detection is irrelevant. As for random movement or a general drift in direction (Sections 6.1 and 6.2), reactive movement is not a problem in aerial survey because of the speed of the platform. However, it could cause a significant bias in the estimation of esw from ship surveys.

Consider the case where whales (or pods) are attracted towards the ship as it approaches and before they are detected by observers. Radial distance, $r$, will be biased downwards as will perpendicular distance, y. Thus, the distributions of $r$ and $y$ will be biased towards zero and estimates of $f(0)$ will be biased upwards causing an overestimate in abundance. The opposite occurs if movement is away from the ship.

Although some models for $f(y)$ are relatively robust to the distortions resulting from vessel avoidance (Burnham et al., 1980), there is in general no satisfactory way to deal with data subject to attraction or avoidance before detection. This is because the changes in $y$ cause the assumption that a whale is equally likely to be present at any y in the vicinity of the platform (Sections 2.1) to be violated. That is, the distribution of whales $G_{i}(y)$ is no longer uniform in the interval ( $0, \mathrm{w}$ ). The observed data are thus a product of the distribution of whales present (after moving) and the distribution of detection probabilities, both of which are unknown. Burnham et al. (1980, pp. 120-25) and Smith (1979) have derived similar expressions to show the effects of reactive movement. Their results are essentially that esw can only be estimated with additional information on the amount of movement.
There have been attempts to investigate reactive movements of dolphins to survey vessels using helicopters. Au and Perryman (1982) found that schools of spotted and spinner dolphins did respond to an approaching ship by moving away from it. However, they did not attempt to ascertain whether or not this reaction occurred before they
were detected from the ship. Hewitt (1985) did collect this information for the same species so that it was possible to calculate the change in perpendicular distance between the time the school was first seen from the air and the time it was detected from the ship in normal searching mode. The amount of movement was insufficient to warrant correction of the observed data, however. Reactive movements of Dall's porpoise have also been investigated in a similar way with mixed results (Bouchet, pers. comm.).

The reaction of Southern Hemisphere minke whales to a ship (an ice-breaker) has also been investigated using a helicopter (Leatherwood, Awbrey and Thomas, 1982). No reaction to the ship was found when it was transiting, although minkes did approach when the ship was stationary.
Data from the 1984/85 IDCR cruise were analysed during the Workshop on Minke Whale Sightings (IWC, 1986a), to look for evidence of vessel avoidance. Some features of the results of the parallel ship experiments indicated that minke whales may have been avoiding the vessels, also estimated values at first sighting were compared with the $y$ estimates when the sighting came abeam: the abeam value was higher in 22 out of the 24 cases. However, the report concluded that 'although the overall results were consistent with vessel avoidance, alternative explanations could be found for all differences that were observed'.

Turnock and Quinn (Turnock, 1987) have applied some new methods for decomposing movement from sighting function which may be relevant in some situations.

## 7. SURVEY DESIGN

### 7.1 Introduction

In order for whale sightings surveys to provide unbiased estimates of population size in an area, it is necessary that the survey cover in some sense a representative part of the area for which a population estimate is desired.

The same principles apply whether the search is a conventional line transect survey, or any of the other forms of survey described in this report, although the details of the analysis are outwardly different. For simplicity, it shall be assumed in the following analysis that a conventional line transect survey is envisaged. It shall also be assumed that the whales are effectively stationary throughout the area to be surveyed: as discussed in Section 6.2, violation of this assumption can lead to serious biases in population estimation for some types of survey design.

### 7.2 Simple random survey design

Conventionally, surveys with the potential to provide unbiased population estimates have been designed by dividing the area to be surveyed into fixed strata, and then randomising the allocation of cruise tracks within each stratum so that each point in the stratum has an equal probability of being covered.
Let $L_{i}$ denote the amount of survey effort (total track length) allocated to stratum $i, A_{i}$ denote the area of stratum i, esw ${ }_{i}$ denote the effective search width for stratum $i$ (which may include an estimate of $g(0)$ ), and $n_{i}$ be
the number of animals sighted. If the track length, $L_{i}$, is randomly allocated over the area, such that each point in the area has an equal probability of being covered, then an unbiased estimate of total population size is:

$$
\hat{\mathrm{N}}=\sum_{\mathrm{i}} \mathrm{n}_{\mathrm{i}} \mathrm{~A}_{\mathrm{i}} /\left(2 \mathrm{~L}_{\mathrm{i}} \mathrm{esw}_{\mathrm{i}}\right)
$$

where the summation is over the number of strata. The coverage probability for each point in stratum $i$ is $2 \mathrm{~L}_{i} \mathrm{esw}_{\mathrm{i}} / \mathrm{A}_{\mathrm{i}}$, which is equal to the effective proportion of the stratum covered by the survey. The formula for the population estimator can be rewritten in the following more general form:

$$
\begin{equation*}
\hat{\mathbf{N}}=\sum_{\mathrm{j}} 1 / \mathrm{p}_{\mathrm{j}} \tag{7.1}
\end{equation*}
$$

where the summation is over all sightings in all strata. $p_{j}$ is the coverage probability for the point at which the jth sighting is obtained. In this simple case, the value of $p_{j}$ is constant within each stratum, but differs between sightings in different strata.


Fig. 10. An example of a saw-tooth survey design.

Randomisation of survey design in order to achieve equal coverage can be achieved in a variety of ways. Fig. 10 gives an example where a rectangular-shaped area is surveyed using a sawtooth survey pattern. The angle between successive transects, and hence the total track length, is fixed in advance, but the starting point is randomly selected on the left-hand edge of the area. The coverage probability is uniform through the area, but the coverage probabilities of different points are not mutually independent.

### 7.3 Variable coverage probability designs

In the case of shipboard surveys, it is not always possible to design a survey so that coverage probability within a stratum is uniform. An example is the IDCR Southern Hemisphere Minke Whale Assessment Cruises, conducted annually since 1978/79. In these surveys, the density of whales is such that it is important to expend substantial survey effort in the neighbourhood of the ice edge, but the requirements of navigating around irregular and unstable ice edges precludes a survey design providing uniform coverage in this area. Before the 1984/85 survey, no attempt was made to ensure random coverage near the ice edge. From 1984/85 onwards, a survey design based on variable coverage probabilities has been used.

The basis for this method is that formula (7.1) can provide unbiased population estimation even when the coverage probabilities at each sighting, $p_{j}$, are non-uniform within a stratum. All that is required is that they can be
calculated, and that the coverage probability is non-zero throughout the area for which a population estimate is required. If $p$ is zero for part of the area, formula (7.1) provides a negatively biased estimator for the total population size, and an unbiased estimator for the population in that part of the area for which $p$ is non-zero.
In calculating the coverage probabilities, it is convenient to divide the coverage probability into the two components:

$$
\begin{equation*}
\mathrm{p}_{\mathrm{j}}=\mathrm{esw}_{\mathrm{j}} \mathrm{p}_{\mathrm{j}}^{\prime} \tag{7.2}
\end{equation*}
$$

where $\mathrm{p}^{\prime}$; is the notional coverage probability per unit effective width, and $e s w_{j}$ is the effective search width at the location of the jth sighting. On the assumption that the search width is small compared with the dimensions of the survey design, $\mathrm{p}^{\prime}{ }_{\mathrm{j}}$ depends only on the survey design (i.e. the process by which the cruise track layout is decided). esw $\mathrm{j}_{\mathrm{j}}$ can be estimated by conventional line transect techniques (i.e. by fitting to perpendicular distance distributions, possibly supplemented by a $g(0)$ estimate from independent observer experiments or other methods - see Sections 4.2, 5).

A practical example of the calculation of the $\mathrm{p}^{\prime}{ }_{\mathrm{j}}$ is given for the 1984/85 IDCR minke whale survey in Antarctic Area IVW by Cooke (1987b). Calculation of coverage probabilities was based on the geometric and other rules by which the cruise tracks were designated. It is not possible to provide general rules for the calculation of $p^{\prime}{ }_{j}$, because the variety of possible survey designs is open-ended.

Use of an average value for esw for part or all of a survey in formula (7.2) will not introduce any bias provided that the $e^{2} w_{j}$ and the $\mathrm{p}_{\mathrm{j}}^{\prime}$ are not correlated. Since it is not possible to evaluate $e_{s w}$ for a single sighting, direct examination of the correlation between $\mathrm{p}_{\mathrm{j}}^{\prime}$ and esw $_{\mathrm{j}}$ is not possible. The approach used by Cooke (1987) is to examine the correlation between $\mathrm{p}^{\prime}{ }_{\mathrm{j}}$ and the perpendicular distance, $\mathrm{y}_{\mathrm{j}}$, using the latter as a proxy for $\mathrm{esw}_{\mathrm{j}}$. If a correlation between $y_{j}$ and $p_{j}^{\prime}$ is indicated, then the bias that would be introduced by using a simple average value of esw can be avoided by weighting each of the data points by $1 / p_{j}^{\prime}$ in the estimation of esw.

At the present time, most commonly available computer programmes for fitting perpendicular distance distributions (such as TRANSECT: Burnham et al., 1980) do not allow weighting of points in this way, although this would be possible in principle. In terms of the point estimate, weighting of points is equivalent to multiple repetition of points in proportion to their weights, but this does not apply to the variance estimates. The problem is analogous to the problem of weighting by school size, discussed in Section 4.3.

### 7.4 Variance estimation

The variance of the estimator is:

$$
\operatorname{var}(\hat{\mathbf{N}})=\mathrm{E}\left(\hat{\mathbf{N}}^{2}\right)-[\mathrm{E}(\hat{\mathbf{N}})]^{2}=\mathrm{E}\left(\hat{\mathbf{N}}^{2}\right)-\mathrm{N}^{2}
$$

where $E$ denotes expectation. Since $\hat{N}^{2}$ itself is an unbiased estimator for $E\left(\hat{N}^{2}\right)$, unbiased estimation of the variance of $\hat{\mathrm{N}}$ requires an unbiased estimator for $\mathrm{N}^{2}$. $\mathrm{N}^{2}$ is the number of pairs of individuals in the population. Hence an unbiased estimator for $\mathrm{N}^{2}$ is given by:

$$
\hat{\mathbf{N}}^{2}=\sum_{i=1}^{n} \sum_{j=1}^{n} 1 / p_{i j}
$$

where $n$ is the number of sightings and $p_{i j}$ is the probability that the points of the ith and jth sightings are both covered by the survey. Hence, unbiased estimation of the variance
of N is possible provided that the pairwise coverage probability is non-zero and calculable for every pair of points in the area. It can also be shown that this is a necessary condition. The estimator for the variance of N is:

$$
\begin{equation*}
\operatorname{var}(\hat{\mathrm{N}})_{\mathrm{EST}}=\sum_{\mathrm{i}=1}^{\mathrm{n}} \sum_{\mathrm{j}=1}^{\mathrm{n}}\left[1 /\left(\mathrm{p}_{\mathrm{i}} \mathrm{p}_{\mathrm{j}}\right)-1 / \mathrm{p}_{\mathrm{ij}}\right] \tag{7.3}
\end{equation*}
$$

If the joint coverage probability is zero for some pairs of positions, then this formula yields a positively biased estimate of the variance.

While unfamiliar in appearance, this variance estimator is nevertheless often identical to more familiar formulae in cases where the latter are valid. For example, suppose that the survey consists of $m$ replicated parallel random transects across a rectangular area of width W and in Fig. 11. Each transect is of fixed width $w$ and is chosen by selecting one of the $\mathrm{W} / \mathrm{w}$ possible transect positions at random. If the area to be surveyed is of width W , and on the assumption that $w$ is small compared with W , then $\mathrm{p}_{\mathrm{i}}=\mathrm{mw} / \mathrm{W}$ everywhere. $\mathrm{p}_{\mathrm{ij}}=\mathrm{p}_{\mathrm{i}}=\mathrm{mw} / \mathrm{W}$ for sightings on the same transect. For sightings on different transects, $\mathrm{p}_{\mathrm{ij}}{ }^{-}$ $=p_{i} p_{j}(m-1) / m=m(m-1)(w / W)^{2}$. The factor ( $m-1$ ) arises from the fact that given the transect covering sighting i has been selected, only ( $\mathrm{m}-1$ ) transects remain to cover sighting j .


Fig. 11. An example of a survey design where parallel transects of width $w$ are selected randomly across the width of the survey area, W.

The population estimate is $\mathrm{nW} /(\mathrm{mw})$ where n is the number of sightings. Applying formula (7.3) yields the following expression for its variance

$$
(\mathrm{W} / \mathrm{w})^{2} /[\mathrm{m}(\mathrm{~m}-1)] \sum_{i=1}^{m}\left(\mathrm{n}_{\mathrm{i}}-\mathrm{n} / \mathrm{m}\right)^{2}
$$

where the summation is over the transects and $n_{i}$ is the number of sightings in the ith transect.

Note that the quantity $n_{i} W / w$ can be regarded as the population estimate based on the data from the ith transect alone. Hence this variance formula is identical to the conventional sample variance formula for the mean of $m$ independent, identically distributed estimators. In more general cases, the two variance estimators would not be equal, and the conventional formula would not be applicable.

In most shipboard surveys, the joint coverage probability is zero for most points, even where the simple coverage probability is non-zero everywhere. For example, for the survey design given in Fig. 10, the joint coverage probability is zero for all points which do not lie on the same member of the family of possible cruise tracks.

The approach used by Cooke (1987b) was to regard the transects as having been randomly and independently chosen within each stratum, even though in actuality the choice of the first transect in each stratum determined the positions of all remaining transects. This yields a variance estimator which may in principle be biased. The reason why it is biased in principle is that the possibility cannot be excluded that whale density varies according to a repeating pattern with period equal to the inter-transect distance so that the transects tend either to all fall in high density areas or all fall in low density areas.

Random rather than regular spacing of the transect end points would avoid the problem. In practice, however, it is doubtful whether this is worthwhile, as the biased variance estimates may be adequate for practical purposes.

### 7.5 Factors to consider in the design of surveys

The analysis of the results of surveys based on variable coverage probability designs is more complex than the analysis of surveys based on the more conventional uniform coverage designs. However, the time and effort involved in the analysis is usually small compared with the expense of the carrying out the survey, so this is not a major consideration. Variable coverage probability designs may enable much greater flexibility in the design of surveys. This is especially important when the survey is influenced by factors not known at the design stage, such as the position of the ice edge, the available time of good weather, etc.

For example, the survey design analysed by Cooke (1987b) enabled the cruise tracks in the neighbourhood of the ice edge to be plotted in real time according to the latest ice information. At the design stage, only the rules for plotting cruise tracks, not the actual tracks themselves, were fixed. That design also enabled some flexibility with regard to the total survey time available - always an unpredictable quantity for whale surveys because survey can only be conducted in sufficiently good weather, although it appeared that the degree of flexibility provided was not sufficient (Anon., 1985)
There is a great variety of ways in which flexibility for weather conditions can be incorporated into the design of the survey, of which the following is just one example: a basic survey for the area is designed which would normally be expected only to use half or less of the available survey time. This is covered first, with pauses during periods of bad weather. The area is then surveyed again according to another design, but this time sections where bad weather is encountered are skipped. Since the probability of encountering bad weather at any point is not a calculable quantity without introducing new and irrelevant assumptions, the coverage probabilities are calculated conditionally on the weather actually experienced. The first part of the survey ensures that coverage is everywhere non-zero.
Problems of irregular ice edges and coastlines apply mainly to shipboard surveys, but problems of bad weather apply to both aerial and shipboard surveys.

### 7.6 Allocation of survey effort

The problem of optimal design of sightings surveys can be expressed as the problem of finding a design which yields the minimum variance of the population estimate, subject to the constraint that it be unbiased, and subject to the constraints of the available search effort.

No general result has been derived for the optimal design of a survey, mainly because of the difficulty of describing the constraints on the available search effort in a general way.

An important special case which can be solved is where the design is restricted to choosing random transects from a fixed set of allowed transects. The question is what selection probability to apply to each of the possible transects. This problem may be relevant for some aerial surveys. The optimal solution can be shown to be that coverage probability should be allocated approximately in proportion to the root mean square of the prior expectation of the density of whales. This means that coverage should be approximately proportional to the prior expectation of density, but with some extra coverage to areas where this prior expectation is very uncertain or where density is liable to be very non-uniform.

This result does not apply if the set of transects is not fixed. However, it appears that optimal allocation is achieved by placing a relatively larger number of short transects in higher density areas and a relatively smaller number of longer transects in lower density areas.

### 7.7 Non-random survey designs

Some whale sightings surveys to date have not been designed according to systematic or random principles. Population estimates have been derived by assuming that the surveys were approximately random designs in some sense. One method of analysing such data is to post-stratify using stratum boundaries which provide reasonably systematic coverage within each stratum (e.g. Hammond, 1986b). This method may not always be possible if no stratum boundaries satisfying this criterion exist. In some cases a wide range of different population estimates can be obtained from different choices of stratification, each of which are individually reasonable. An analogous problem occurs in politics, where the result of an election may be as much dependent on the choice of electoral district boundaries as on the overall proportions of votes in the population.

An alternative approach is contouring (Best and Butterworth, 1980). However, contouring as a method has been rejected by the IWC Scientific Committee (IWC, 1984), not necessarily because it is seriously biased relative to other methods, but because the bias in unquantifiable, and because valid variance estimates cannot easily be obtained. It also relies on an arbitrary choice of aggregation units for point density estimates, and on an arbitrary choice of interpolation method.

Estimates obtained from non-random surveys are pron to positive bias if the areas actually covered have been chosen because they were known or suspected to contain high whale densities, and if the method of post-stratification does not or cannot take account of this.

Even if large-scale choice of survey tracks are reasonably representative, biases can arise if the searching vessel tends to leave its planned cruise track and make small excursions when whales are encountered, if the distribution of whales is patchy. This can be a feature of whalès sightings surveys which have supplementary duties such as whale marking.

Methods of post-stratification can be devised to reduce this bias, but it is likely that more than one method would be available, yielding different results.

The judgement of whether population estimates obtained from past surveys which lacked a random design are sufficiently reliable for use in whale population assessments has to be made on a case by case basis.

## 8. CUE COUNTING AND ESTIMATION OF SURFACING RATE

### 8.1 Estimation of abundance from cue counting data

In Section 2.5 the statistic $\mathrm{n} /(\mathrm{BTesa}$ ) was suggested as an estimator of local density, where $n$ represents the number of surfacing cues counted during $T$ hours on effort, $B$ the mean surfacing rate - surfacing per whale per hour - for the whales in that area, and esa the effective search area estimated from the distances to the cues recorded and (possibly) the proportion of 'duplicate' cues. How can this method be applied to estimation of abundance and what is the appropriate measure of B ?

It is possible to obtain answers to both questions by using an approach analogous to that used by Cooke (1984) for line transect sampling (see also Section 7.3). In that approach a random variable $x_{i}$ is defined for each whale in the population, which takes the value 1 for detected whales and 0 for undetected whales, with expectation $p_{i}$ which can be deduced from the design of the survey. Thus $x_{i} / p_{i}$ has the expectation 1 for each of the $\mathbf{N}$ whales in the population, and the statistic

$$
\sum_{i=1}^{N} x_{i} / p_{i}
$$

has the expectation $N$, that is, it is unbiased for $N$.
In attempting to derive an unbiased estimator for N from cue counting data we replace the dummy variable $\mathrm{x}_{\mathrm{i}}$ with the number of cues counted from the ith whale, say $x_{i}{ }_{i}$. We then require the expectation of $\mathrm{x}_{\mathrm{i}}{ }_{\mathrm{i}}$. To derive $\mathrm{E}\left(\mathrm{x}^{\prime}{ }_{i}\right)$ it is convenient to envisage the esa as a 'hard-edged' region of arbitrary shape having an average length $\mathrm{h}(\mathrm{y})$ parallel to the trackline at perpendicular distance y. As before let $G_{i}(y)$ represent the pdf for perpendicular distance of the ith whale from the trackline. A whale at distance $y$ from the trackline spends a period $h(y) / v_{i}$ within the esa where $v_{i}$ is the speed of the platform as it passes whale i. Thus the expected number of cues counted from whale $i$ is

$$
\int_{0}^{\infty} G_{i}(y)\left[h(y) / V_{i}\right] B_{i} d y
$$

where $B_{i}$ is the surfacing rate parameter appropriate to whale $i$ at the time of the survey. Assuming $G_{i}(y)$ is constant at $p_{i}$ over the range of $y$ for which $h(y)$ is greater than zero,

$$
\mathrm{E}\left(\mathrm{x}_{\mathrm{i}}^{\prime}\right)=\left(\mathrm{p}^{\prime}{ }_{\mathrm{i}} \mathrm{~B}_{\mathrm{i}} / \mathrm{V}_{\mathrm{i}}\right) \int_{0}^{\infty} \mathrm{h}(\mathrm{y}) \mathrm{dy}=\mathrm{p}^{\prime}{ }_{\mathrm{i}} \mathrm{~B}_{\mathrm{i}} \mathrm{esa} / 2 V_{\mathrm{i}}
$$

On this basis, as unbiased estimator for $N$, given $B_{i}$, is

$$
\sum_{i=1}^{N} x^{\prime}{ }_{i} 2 V_{i} /\left(p^{\prime}{ }_{i} B_{i} e s a\right) .
$$

There are two approaches possible with respect to the required $B_{i}$ values. One is to incorporate an estimate of $B_{i}$ for each whale detected; the other is to use a single value for surfacing rate, so that

$$
\mathrm{N}=(1 / \mathrm{B}) \sum_{\mathrm{i}=1}^{\mathrm{N}} \mathrm{x}^{\prime} 2 \mathrm{~V}_{\mathrm{i}} /\left(\mathrm{p}^{\prime}{ }_{\mathrm{i}} \mathrm{esa}\right)
$$

Equating the expectation of this estimator with N identifies the required value of $B$ as the mean of $B_{i}$ for all $N$ whales in the population. One possibility is to use the same nominal value of $B$ for all surveys of the chosen area on the basis that although this may give only a rough indication of total abundance it will provide an effective index. That strategy is reasonable only if the mean of $B_{i}$ is very similar on each survey. The alternative is to attempt to estimate the $1 / B_{i}$ values at the time of the survey. Section 8.2 considers the estimation of 'blow rates' using observations from a survey vessel. For shipboard survey it would be possible to carry out such experiments on a representative sample of whales, for example, estimate the blow rate from every kth pod detected and take the weighted mean

$$
\left(\sum_{j=1}^{n / k} k \theta_{j} / B_{j}\right) / \sum_{i=1}^{n} \theta_{i}
$$

where $\theta_{\mathrm{i}}$ represents the weighting $\mathrm{x}^{\prime}{ }_{\mathrm{i}} \mathrm{v}_{\mathrm{i}} / \mathrm{p}^{\prime}{ }_{\mathbf{i}}$.
Certain biases are inherent in that approach, for example, using every kth pod biases the blow rate estimation in favour of whales in small pods, whereas the difficulty of successfully completing estimation experiments on small pods biases estimation in favour of whales in large pods.

For aerial survey, extended blow rate estimation experiments are clearly impractical and we therefore seek unbiased estimates of $1 / \mathrm{B}_{\mathrm{i}}$ appropriate for that situation. $1 / B_{i}$ is the average inter-surfacing interval for the ith whale at the time of the survey ('inter-surface intervals' and 'dive time' could be used synonymously; however, 'dive' might be taken to imply travel to a certain minimum depth), thus it is natural to consider the average observed inter-surface interval for a detected whale as a candidate. That is, if $\mathrm{X}^{\prime}{ }_{i}$ surface cues are counted from the ith whale within the sector, measure the time from the first surfacing counted to the first surfacing subsequent to the whale leaving the sector, say $t_{i}$, and use $t_{i} / x_{i}$. Replacing $1 / B_{i}$ by $t_{i} / x_{i}^{\prime}$ in the estimate gives

$$
\hat{N}=\sum_{i=1}^{N} t_{i} 2 V_{i} /\left(p^{\prime}{ }_{i} e s a\right)
$$

This suggests, then, that for estimation of N (though not esa) the surface counts $x_{i}^{\prime}$ are irrelevant if the time $t_{i}$ can be measured ( $\mathrm{t}_{\mathrm{i}}=0$ for undetected whales). The basis for this estimator is that for each whale, irrespective of its rate or pattern of surfacing, the expected time between its first surfacing within the sector and its first surfacing 'behind' the sector is the same, i.e. $\mathrm{p}^{\prime}{ }_{\mathrm{i}} \mathrm{esa} / \mathrm{v}_{\mathrm{i}}$. The derivation of the estimator was based on the concept of a 'hard-edged' region for esa; however, it is straightforward to show it is unbiased for N in the realistic situation where the cue counting sector has no leading edge, if each whale can surface at most once within visible range of the platform. In the general case the properties of this estimator have, as yet, been checked using simulation trials only.

The use of the $t_{i}$ measure would seem to provide a very powerful technique for abundance estimation, because of the independence of $E\left(t_{i}\right)$ and the surfacing rate. It is therefore worth considering the practical aspects. For aerial survey the basic cue counting procedure (see Sections 2.5 and 3.1 ) would be extended to record, from each whale detected in the sector, the time of its first surfacing behind the abeam line. That information would form part of the data collected during confirmation. We can then envisage three basic problems: failure to relocate
the detected whale behind the abeam line; confusion of the detected whale with an undetected member of the same pod; and loss of data from a second or third sighting still within the sector after the first has passed abeam. In case the first problem is significant it may be necessary to substitute nominal values of $t_{i}$ for measured values. For example, it may be that inter-surface intervals are of two main types, those corresponding to long dives and those corresponding to a breathing sequence with the whale near the surface throughout. In that case the main source of variation in surfacing rate may be variation in the proportion of long dives. If failure to relocate a whale was due to its having started a long dive subsequent to its last surfacing then an average time for the long dive category could be added to the time of the last surfacing to give $t_{i}$. In some cases the type of dive may be apparent from the nature of the surfacing, e.g. 'fluking' might be associated with the start of a long dive. The second problem could arise if, for example, two whales from a pod of three are detected in a sector, the third surfacing behind the abeam line before one or both of the two detected whales, so that sum of the two $t_{i}$ values is biased downwards. If this problem is significant the method would be applied to pods rather than whales, thus $t_{i}$ is the time between the first surface cue from the pod within the sector and the first behind the sector. This would certainly be the appropriate strategy for shipboard sampling. Note that $E\left(t_{i}\right)$ is still $\mathrm{p}^{\prime}{ }_{\mathrm{i}} \mathrm{esa} / \mathrm{vi}$ even for pods. The estimator of whale abundance is then

$$
\hat{\mathbf{N}}=\sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{t}_{\mathrm{i}} \mathrm{~s}_{\mathrm{i}} 2 \mathrm{v}_{\mathrm{i}} /\left(\mathrm{p}_{\mathrm{i}}^{\prime} \mathrm{esa}\right)
$$

where $s_{i}$ is the estimated pod size.
It is impossible to estimate $t_{i}$ for all detected whales or pods if more than one may be in the sector simultaneously. In this case we might confirm only every kth target, assuming the probability of more than k targets within the sector simultaneously is negligible. Then

and $\theta_{i}$ is the weighting $\mathrm{v}_{\mathrm{i}} / \mathrm{p}_{\mathrm{i}}^{\prime}$.
The properties of this type of estimator require further evaluation. In the meantime, it is recommended that $t_{i}$ data be collected where convenient.

### 8.2 Estimation of surfacing rate

Within the context of cetacean survey the objectives of measuring surfacing rates are two-fold: to determine the appropriate values for estimators such as those suggested in the last section; and to determine what factors most affect the rate and may therefore confound comparisons between surveys. Reliable information on surfacing rates is one of the most immediate results of short term radio or acoustic tagging and a number of studies have provided such results, e.g. Watkins, Moore, Wartzok and Johnson (1981) for fin whales; Harvey and Mate (1984) for gray whales.

There have been a large number of observational studies of diving behaviour, an extensive review of existing data was prepared by Leatherwood, Goodrich, Kinter and Truppo (1982). Information was obtained from 54 researchers; however, the wide variety of objectives and experimental procedures make comparison of the results


Fig. 12. Results of blow rate experiments carried out on vessel SM2 during the 1986/87 IWC/IDCR Southern Hemisphere minke whale assessment cruise. From Ward (1988). The pod size for each experiment is given after the experiment letter code.
difficult. Observational studies reported since 1982 include those by Winn and Martin (1983) on right whale, fin whale and humpback off the northeast US coast, Würsig, Dorsey, Fraker, Payne, Richardson and Wells (1984) on bowheads in the Beaufort sea, Würsig, Wells and Croll (1986) on gray whales in the Bering Sea, Dolphin (1987) on humpback whales, and Whitehead (1985) and Gordon (1987) for sperm whales. Caldwell, Caldwell and Rice (1966) reviewed observations on sperm whale diving and respiration patterns from whales. The humpback data from Dolphin (1987) is very detailed and demonstrates effects of different behaviour on variables of the dive cycle such as dive time, surface time and number of blows on surfacing. Blow rate over the complete dive cycle was less affected by behaviour than the component variables, with extremes of 0.7 blows per minute and 1.3 blows per minute corresponding to 'fast travel' and 'surface display'. A
further review of available data would seem worthwhile in view of the recent increase in observational and tagging studies.
Information on minke whale surfacing rates has been obtained by observation from the IDCR Antarctic Cruises and reported in Joyce (1982), Hiby and Ward (1986b) and Ward and Hiby (1987). The main problem encountered in such studies is that of ensuring that all surfacings from the pod under observation are recorded during the chosen observation period. The experimental procedures used during the IDCR cruises were designed to minimise bias due to losing pods during long dives. By recording the exact time of each blow seen from the observed pod (Ward, 1988) using tape recording equipment described in Section 3.1 was able to identify experiments in which some surfacings had not been recorded. Fig. 12 shows the times at which blows were recorded from 7 different pods. The
blow rates in cases A, C, E, F and G vary from 29 to 46 blows per whale per hour. In case B and D the rate was only 16 per whale per hour; however, it is evident from the timing of recorded blows for the two animals in case D that a complete surfacing sequence was missed during the observation period. The average rate for the other pods was 34 blows per whale per hour. The average values obtained for the studies in $84 / 85$ and $85 / 86$ were 35 and 33 . The average for the study by Joyce was 37 blows per whale per hour.

### 8.3 Analysis of cue counting data from independent observers

Since the first presentation of this review, two surveys for minke whales have been carried out using cue counting sampling (Hiby and Ward, 1988; Hiby et al., 1989). These studies have indicated the value of obtaining data from independent observers and it is therefore important to include a brief description of the methods used in this revision.
8.3.1 Data collection and analysis using the 'product rule' Sections 2.5 and 3.4 introduced the ideas of recording cues from a defined area of sea surface, and identifying 'duplicate' cues from independent observers by comparing their times of occurrence. In both the studies referred to above the data were recorded onto stereo tape recorders, as described in section 3.1, and the cues recorded by each set of observers listed in chronological order. If an observer recorded, say, 'two dives......minke whales ......coming abeam now at declination angle 32 degrees' and the time of the dives read from the display was, say, 2:11:42, then $2: 11: 42$ was entered twice in the list for that observer. The two lists were then compared and cues recorded as occurring at precisely the same time identified as duplicates and removed from the lists. The comparison was then repeated to identify cues occurring within 1 second of each other, and so on up to a maximum allowable difference of $t$ seconds. The number of duplicates identified in this way thus depends on the value of $t$ used, however, both studies indicated little or no increase beyond at of 2 or 3 seconds, suggesting this algorithm leads to reliable and objective identification of duplicates.

Section 4.2 contains estimates for $g(0)$ or $g(0) / f(0)$ in closed form, however, the cue counting data have been treated differently, using numerical methods to derive ML estimates of these quantities. The distance estimates for the detected cues were first grouped into M intervals and the resulting frequencies ordered into an array O where $\mathrm{O}(\mathrm{I}, \mathrm{J})$ represents the number of cues seen by the primary observers and allocated by them to the $I^{\text {th }}$ distance interval which were also seen by the secondary observers and allocated by them to the $\mathrm{J}^{\text {th }}$ distance interval. The last column and row were used for the non-duplicates, thus $\mathrm{O}(\mathrm{I}, \mathrm{M}+1)$ represented the number of cues seen by the primary observers and allocated by them to the Ith interval which were not seen by the secondary observers; similarly for $\mathbf{O}(\mathrm{M}+1, \mathrm{~J})$.

The array contains information on the reliability of duplicate identification and distance estimation; given reliable identification of duplicates and accurate distance estimation we would expect entries on the principal diagonal, last row and last column only. Spurious duplicates would appear scattered across the array because distance estimates would be uncorrelated. Variance and bias in the distance estimates leads to spreading of the
diagonal 'ridge' and shifting of the ridge from the diagonal, respectively.

The effective search area is estimated by maximising the likelihood of the data in the array with respect to the parameters of the sighting functions for the two sets of observers. The array frequencies are assumed to follow a multinomial distribution with the probability for position I, J given by

$$
\mathrm{P}(\mathrm{I}, \mathrm{~J})=\int_{0}^{w} \mathrm{~g}_{1}(\mathrm{r}) \mathrm{f}_{1}(\mathrm{I}, \mathrm{r}) \mathrm{g}_{2}(\mathrm{r}) \mathrm{f}_{2}(\mathrm{~J}, \mathrm{r})\left(2 \mathrm{r} / \mathrm{w}^{2}\right) \mathrm{dr}
$$

for $I$ and $J$ from 1 to $M$ (for $J=M+1$ replace $g_{2}(r) f_{2}(J, r)$ by $1-\mathrm{g}_{2}(\mathrm{r})$; similarly for $\mathrm{I}=\mathrm{M}+1$ ).

Here $g_{1}$ and $g_{2}$ represent the sighting functions for the primary and secondary observers; $f_{1}(I, r)$ and $f_{2}(J, r)$ are functions giving the probability the primary observers allocate a cue to interval $I$, and the secondary observers allocate the cue to interval J , given the cue occurs at distance $\mathrm{r} . \mathrm{W}$ is the maximum distance at which cues are detected. Error in distance estimation is thus incorporated into the model for the recorded data, rather than smearing the data to account for error (see Sections 3.1, 4.1.2, 5). The parameters of the $f$ functions can also be estimated by maximising the likelihood; alternatively it may be possible to evaluate these functions using distance estimation experiments. The term $2 \mathrm{r} / \mathrm{w}^{2}$ in the integral is the probability density for the position of a cue with respect to distance from the platform assuming no reaction by the whale to the platform.

### 8.3.2 Replacing the product rule

In the previous section the probability a cue at $r$ is seen by both sets of observers was modelled as the product $\mathrm{g}_{1}(\mathrm{r}) \mathrm{g}_{2}(\mathrm{r})$; as pointed out in Section 4.2.2(f), this assumption of independence may not be valid. One relevant factor in cue counting sampling is that detection of one cue leads to increased probability of detection of further cues from that whale or pod. Indeed, in both studies referenced, observers counted all cues from detected pods, so that the probability of seeing cues from detected pods was equal or close to 1 . This results in the probability of duplicate detection exceeding the product $g_{1}(r) g_{2}(r)$. However, this problem, which is significant only in the case of shipboard sampling, can be dealt with fairly easily by making the probability of a cue being seen conditional on whether or not the pod has been detected previously. The product rule is then assumed to apply only to cues from undetected pods, and the unconditional cue detection probabilities calculated by considering the proportion of pods detected as a function of $r$. The technique requires a probability distribution for the variation in surfacing rate between pods. One possibility is to assume this is proportional to the variation in pod size, which can be modelled using existing data.

Despite this modification, some data sets still show an excess number of duplicate cues observed over expected as $r$ increases. This presents a dilemma, at least for minke whale surveys, because experience suggests that data from independent observers is essential. For example, esa's estimated from aerial surveys of minke whales in three different areas of the North Atlantic (Hiby et al., 1989) differed widely when using the assumption $g(0)=1$, suggesting that observers on some surveys may search at larger distances and thus have reduced $g(0)$ values. Also, data from shipboard surveys of Antarctic minke whales (Hiby and Ward, 1986b; Hiby and Ward, 1988) produced
frequency distributions for distance to cues which show little or no 'shoulder', and for which reliable estimation of esa needs data from independent observers (c.f. Section 4.2.2(a)). This suggests that the product rule may not provide a useful model for duplicate sighting probability.

There is actually no good reason to suppose that it should hold in general. To see this it is useful to consider how detection probability might change with the 'strength' of the cue at the eye. The visibility of, say, a blow is known to vary in response to many factors, such as the size and density of the blow, its duration, the colour of the background, and its distance from the observer. Of these, only distance is dealt with explicitly in the analysis. Let the remaining factors combine to produce a cue with a source strength, c. At the eye of the observer, the strength of the cue, $c_{e}$, depends on its distance from the observer; it may be reasonable to assume that $c_{e}$ reduces as an inverse power of distance, i.e. $c_{e}=c / r^{b}$. It may also be necessary to include a term for light absorption and scattering; thus $c_{e}=c / r^{b} . e^{-\lambda r}$

Fig. 13 shows how the probability of detection, say $\rho$, may change with the strength of the cue at the eye of the observer. Thus, at very low cue strength, $\rho$ is zero and increases with cue strength to an asymptotic value which depends on the searching strategy. For example, for an observer searching without binoculars, so that the entire search area is within the field of view, detection probability may approach 1 if the observer is able to maintain concentration continuously. It is only the conditional probability of detection, $\rho$, which has the potential to be independent between sets of observers: variation in $c_{e}$ will affect both sets.

Duplicate detection probability is the average over $c_{e}$ of the duplicate detection probability conditional on $c_{e}$. If the function is the same for both observers

$$
\begin{aligned}
\mathrm{g}_{12}(\mathrm{r}) & =\mathrm{E}\left(\rho^{2}\right) \\
& =\mathrm{E}^{2}(\rho)+\operatorname{var}(\rho) \\
& =\mathrm{g}^{2}(\mathrm{r})+\operatorname{var}(\rho)
\end{aligned}
$$



Fig. 13. Cue detection probability $\rho$ as a function of cue strength at the eye, $c_{e}$. The continuous line shows a gradual increase in $\rho$ to an asymptotic value which equals $g(0)$, assuming $\operatorname{var}(\rho)$ tends to zero as $r$ tends to zero. In the case of the dotted line, detection probability approximates a step function at threshold strength $c_{e}^{*}$.

If functions, and hence $g(r)$ functions, are different for the two sets, then

$$
g_{12}(r)=g_{1}(r) \cdot g_{2}(r)+\operatorname{covar}\left(\rho_{1} \rho_{2}\right)
$$

Thus the duplicate probability exceeds the product of the individual detection probabilities by the variance of $\rho$ (or covariance of $\rho_{1}$ and $\rho_{2}$ ) resulting from the variation in cue strength.

How, then, can data from independent observers be used, given that $\operatorname{var}(\rho)$ is an unknown function of $r$ ? One way to proceed is to assume that $\operatorname{var}(\rho)$ tends to zero as $r$ tends to zero because the variation in $c_{e}$ then falls under the asympotic section of the $\rho$ function. The product $g_{1}(r)^{e}$ $g_{2}(r)$ in the formula for $P(I, J)$ above can then be replaced by $g_{1 \mid 2}(r) \cdot g_{2}(r)$, where $g_{1 \mid 2}(r)$ is the conditional probability that the main observers detect a cue at r given that the secondary observers detect it, and the parameters of $g_{1 \mid 2}(r)$ estimated by maximising the likelihood of the O array, eliminating the last column (i.e. sightings by the primary observers only). Finally, $\mathrm{g}_{1 \mid 2}(0)$ can be used to estimate $g_{1}(0)$ on the grounds that if, as $r$ tends to zero, $\operatorname{var}(\rho)$ tends to zero, then $g_{12}(r)$ tends to $g_{1}(r) g_{2}(r)$ and $g_{1 \mid 2}(0)$ tends to $\mathrm{g}_{1}(0)$.

An alternative procedure is to adopt functional forms for and the pdf of $c$ and derive $g(r)$, rather than model $g(r)$ directly. Such functions would need to be based on results of research into human visual perception. The advantages of this are that it would make better use of the duplicate data and also reduce the uncertainty about the form of $g(r)$ which currently precludes reliable estimation of effective search range when the data are 'spiked'. For example, it may be valid to assume that the probability of cue detection changes from 0 to $g(0)$ at some threshold strength $c_{e}{ }^{*}$, so that $g(r)$ is simply $g(0)$ times the probability that $c_{e}$ exceeds $c_{e}{ }^{*}$ (cf Fig. 13). In that case var $(\rho)$ equals $g(r)(g(0)-g(r))$ and duplicate detection probability equals $\mathrm{g}(0) \mathrm{g}(\mathrm{r})$. If $\mathrm{g}(\mathrm{r})$ differs between the two observers,

$$
\mathrm{g}_{12}(\mathrm{r})=\mathrm{g}_{1}(0) \mathrm{g}_{2}(0) \min \left[\mathrm{g}_{1}(\mathrm{r}) / \mathrm{g}_{1}(0), \mathrm{g}_{2}(\mathrm{r}) / \mathrm{g}_{2}(0)\right]
$$

This model implies that the expected proportion of the sightings made by one set of observers which are also detected by the other set does not decline as $r$ increases. This is consistent with some survey data. Furthermore, detection at $r$ occurs if $c_{e}=c / r^{b}>c_{e}{ }^{*}$, i.e. $c>r^{b} c_{e}{ }^{*}$, so that if $c$ has pdf $f(c)$,

$$
g(r)=g(0) \int_{r^{b} c_{e}{ }^{*}}^{\infty} f(c) d c
$$

For example, if elements determining cue strength (size, contrast, duration) combine multiplicatively, so that c follows a lognormal distribution, then

$$
\mathrm{g}(\mathrm{r})=\mathrm{g}(0) \int_{\mathrm{r}^{\mathrm{b}} \mathrm{c}_{\mathrm{e}}{ }^{*}}^{\infty}\left[1 /\left(\mathrm{c} \sqrt{ } 2 \pi \sigma^{2}\right)\right] \exp \left[-(\log c-\mu)^{2} / 2 \sigma^{2}\right] \mathrm{dc} .
$$

Changing the variable of integration from $c$ to $x=(1 / b) \log c$ and defining units such that $\mathrm{c}_{\mathrm{e}}{ }^{*}=1$, then

$$
\mathrm{g}(\mathrm{r})=\mathrm{g}(0)\left\{1-\Phi\left(\left(\log \mathrm{r}-\mu_{1}\right) / \sigma_{1}\right)\right\}
$$

where $\mu_{1}=\mu / \mathrm{b}$ and $\sigma_{1}=\sigma / b$ and $\Phi$ denotes the unit normal integral.
The effective search area would then be estimated by maximising the likelihood of the observed data in the $O(I, J)$ array with respect to the parameters $\mu_{1}$ and $\sigma_{1}$.

## ACKNOWLEDGEMENTS

This report has tried to cover a large amount of material and would not have been possible without contributions from Justin Cooke (Sections 4.2 and 7) and Steve Buckland (Section 5). We are also indebted to Tony Martin and Tim Waters who compiled an Appendix to the original report (SC/39/0 8) which is not included in this publication. These contributions have been edited by us and final responsibility for their content rests with us. A special acknowledgement is due to Justin Cooke for valuable contributions to several of the topics covered by the report, to Terry Quinn for helpful review comments, to Gilly Banks for the preparation of the manuscript under very difficult conditions and to Helen Coulson for coping with our failure to keep to the IWC style guide. Finally thanks to Greg Donovan for patience!

## REFERENCES

Alling, A., Gordon, J., Rice, M. and Whitehead, H. 1983. WWF Netherlands Indian Ocean Sperm Whale Study. 1983 Interim Report. Report to WWF/IUCN. 29pp.
Anonymous. 1985. Report of the 1984/85 IWC/IDCR Southern Hemisphere minke whale assessment cruise, Area IV. Paper SC/J85/S11 presented to the IWC Scientific Committee Workshop on Minke Whale Sightings, June 1987. (Unpublished)
Anonymous. 1987. Report of the $1986 / 87$ IWC/IDCR Southern Hemisphere minke whale assessement cruise. Paper SC/39/Mi14 presented to the IWC Scientific Committee, June 1987 (Unpublished).
Au, D. and Perryman, W. 1982. Movement and speed of dolphin schools responding to an approaching ship. Fish. Bull. 80(2):371-9
Au, D. and Perryman, W. 1985. Dolphin habitats in the Eastern Tropical Pacific. Fish. Bull. 83(4):623-44.
Basson, M. and Butterworth, D.S. 1984. An exploratory investigation of the use of a hazard-rate model to assess the target motion correction factor in sighting estimates of minke whale abundance. Rep. int. Whal. Commn 34:309-14.
Best, P.B. 1985. The recovery of right whales off South Africa. Paper SC/37/PS28 presented to the IWC Scientific Committee, June 1985. (Unpublished).
Best, P.B. and Butterworth, D.S. 1980. Report of the Southern Hemisphere Minke Whale Assessment Cruise, 1978/79. Rep. int. Whal. Commn 30:257-83.
Buckland, S.T. 1985. Perpendicular distance models for line transect sampling. Biometrics 41:177-95.
Buckland, S.T. 1986. Estimating minke whale density - a suggested procedure. Annex H of the Report of the Workshop on Minke Whale Sightings. Rep. int. Whal. Commn 36:469-70
Buckland, S.T. 1987. An assessment of the performance of line transect models for fitting IWC/IDCR cruise data, 1978/79 to 1984/85. Rep. int. Whal. Commn 37: 277-9.
Buckland, S.T. and Anganuzzi, A.A. 1988. Comparison of smearing methods in the analysis of minke sightings data from IWC/IDCR Antarctic cruises. Rep. int. Whal. Commn 38: 257-63.
Burnham, K.P. and Anderson, D.R. 1984. The need for distance data in transect counts. J. Wildl. Manage. 48(4): 1,248-54.
Burnham, K.P., Anderson, D.R. and Laake, J.L. 1980. Estimation of density from line transect sampling of biological populations. Wildlife Monograph No. 72:1-202.
Butterworth, D.S. 1982a. On the functional form used for $g(y)$ for minke whale sightings and bias in its estimation due to measurement inaccuracies. Rep. int. Whal. Commn 32:883-88.
Butterworth, D.S. 1982b. A possible basis for choosing a functional form for the distribution of sightings with right angle distance: some preliminary ideas. Rep. int. Whal. Commn 32:555-61.
Butterworth, D.S., Best, P.B. and Basson, M. 1982. Results of analysis of sighting experiments carried out during the 1980/81 Southern Hemisphere Minke Whale Assessment Cruise. Rep. int. Whal. Commn 32:819-34.
Butterworth, D.S., Best, P.B. and Hembree, D. 1984. Analysis of experiments carried out during the 1981/82 IWC/IDCR Antarctic minke whale assessment cruise in Area II. Rep. int. Whal. Commn 34:365-92.

Butterworth, D.S. and Borchers, D.L. 1988. Estimates of $g(0)$ for minke schools from the results of the independent observer expcriment on the 1985/86 and 1986/87 IWC/IDCR Antarctic assessment cruises. Rep. int. Whal. Commn 38: 301-13.
Butterworth, D.S. and McQuaid, L.H. 1986. An initial analysis of experiments carried out on the 1984/85 IWC/IDCR Antarctic minke whale assessment cruise to compare closing and passing mode procedures in respect of minke whale density estimation. Paper SC/38/Mi13 presented to the IWC Scientific Committee, June 1986. (Unpublished).
Caldwell, D.K., Caldwell, M.C. and Rice, D.W. 1966. Behaviour of the sperm whale Physeter catodon L. pp. 755-89. In: K.S. Norris (ed.) Whales, Dolphins and Porpoises. Univ. Calif. Press, Berkeley.
Clarke, R. 1982. An index of sighting conditions for surveys of whales and dolphins. Rep. int. Whal. Commn 32:559-61
Cooke, J.G. 1984. The estimation of whale populations of non-uniform density from sightings, with implications for survey design. Paper SC/36/Mi29 presented to the IWC Scientific Committee, May 1984. (Unpublished).
Cooke, J.G. 1985. Notes on the estimation of whale density from line transects. Rep. int. Whal. Commn 35:319-24.
Cooke, J.G. 1986a. On the net recruitment rate of gray whales with reference to interspecific comparisons. Rep. int. Whal. Commn 36: 363-6.
Cooke, J.G. 1986b. Further notes on the estimation of whale density from shipborne line transects. Paper $\mathrm{SC} / 38 / \mathrm{Mi} 20$ presented to the IWC Scientific Committee, June 1986. (Unpublished).
Cooke, J.G. 1987a. Further analyses of parallel ship experiments conducted during the 1984/85 IDCR cruise. Paper SC/39/Mi22 presented to the IWC Scientific Committee, June 1987.
Cooke, J.G. 1987b. Estimation of the population of minke whales in Antarctic Area IVW in 1984/85. Rep. int. Whal. Commn 37: 273-6.
Dohl, T.P., Bonnell, M.L. and Ford, R.G. 1986. Distribution and abundance of common dolphin Delphinus delphis in the southern California USA blight, a quantitative assessment based on aerial transect data. Fish. Bull. 84(2):333-44.
Doi, T., Kasamatsu, F. and Nakano, T. 1982. A simulation study on sighting survey of minke whales in the Antarctic. Rep. int. Whal. Commn 32:919-28.
Dolphin, W.F. 1987. Dive behaviour and estimated energy expenditure of foraging humpback whales in southeast Alaska. Can. J. Zool. 65:354-62.
Donovan, G.P. and Gunnlaugsson, Th. 1989. North Atlantic Sightings Survey 1987: report of the aerial survey off Iceland. Rep. int. Whal. Commn 39:437-41.
Drummer, T.D. and McDonald, L.L. 1987. Size bias in line transect sampling. Biometrics 43: 13-21.
Ellison, W.T., Clark, C.W. and Beeman, K. 1985. Acoustic location techniques and calibration methods used during the Spring 1984 and 1985 bowhead whale, Balaena mysticetus, migrations. Paper SC/37/PS10 presented to the IWC Scientific Committee, June 1985 (unpublished) 22pp.
Gates, C.E. 1969. Simulation study of estimation for the line transect sampling method. Biometrics 25:317-28.
Gentleman, R. and Zeh, J.E. 1987. A statistical model for estimating the number of bowhead whales, Balaena mysticetus, passing a census point from combined visual and acoustic data. Rep. int. Whal. Commn 37:313-27.
Gordon, J.C.D. 1987. Behaviour and ecology of sperm whales off Sri Lanka. Unpub. PhD thesis. University of Cambridge, England.
Gunnlaugsson, T. and Sigurjónsson, J. 1988. NASS-87: Estimation of abundance of large cetaceans from observations made on board Icelandic and Faroese survey vessels. Paper SC/40/O 30 presented to the IWC Scientific Committee, May 1988 (unpublished).
Hammond, P.S. 1984a. On the application of line transect sampling to the estimation of the number of bowhead whales passing the Point Barrow ice-camps. Rep. int. Whal. Commn 34:465-7.
Hammond, P.S. 1984b. An investigation into the effects of different techniques of smearing the IWC/IDCR minke whale sightings data and the use of different models to estimate density of schools. Rep. int. Whal. Commn 34:301-9.
Hammond, P.S. 1986a. Line transect sampling of dolphin populations. pp. 251-79. In: M.M. Bryden and R. Harrison (eds.) Research on Dolphins. Oxford University Press.
Hammond, P.S. 1986b. On the post-stratification of sightings data from the 1978/79 - 1982/83 IWC/IDCR Southern Hemisphere minke whale assessment cruises and survey design for future cruises. Rep. int. Whal. Commn 36:225-38.

Hammond, P.S. and Laake, J.L. 1983. Trends in estimates of abundance of dolphins involved in the purse-seining fishery for tunas in the eastern tropical Pacific Ocean, 1977-81. Rep. int. Whal. Commn 33:565-88.
Harvey, J.T. and Mate, B.R. 1984. Dive characteristics and movements of radio tagged gray whales in San Ignacio Lagoon, Baja California Sur, Mexico. pp. 561-75. In: M.L. Jones, S.L. Swartz and S. Leatherwood (eds). The Gray Whale. Academic Press.
Hay, K. 1982. Aerial line transect estimates of abundance of humpback, fin and long-finned pilot whales in the Newfoundland-Labrador area. Rep. int. Whal. Commn 32:475-86.
Hayes, R.J. and Buckland, S.T. 1983. Radial distance models for the line transect method. Biometrics 39:29-42.
Herzing, D.L. and Mate, B.R. 1984. Gray whale migrations along the Oregon coast, 1978-1981. pp. 289-307. In: M.L. Jones, S.L. Swartz and S. Leatherwood (eds). The Gray Whale. Academic Press.
Hewitt, R.P. 1985. Reactions of dolphins to a survey vessel: effects on census data. Fish. Bull. 83:187-93.
Hiby, A.R. 1982. The effect of random whale movement on density estimates obtained from whale sighting surveys. Rep. int. Whal. Commn 32:791-4.
Hiby, A.R. 1985. An approach to estimating population densities of great whales from sighting surveys. IMA Journal of Mathematics Applied in Medicine and Biology 2:201-20.
Hiby, A.R. 1987. Census techniques for small populations with special reference to the Mediterranean monk seal. Symp. Zool. Soc. Lond.
Hiby, A.R., Martin, A.R. and Fairfield, F. 1984. IDCR cruise/aerial survey in the North Atlantic 1982: aerial survey report. Rep. int. Whal. Commn 34:633-44.
Hiby, A.R. and Thompson, D. 1985. An analysis of sightings data from the 1983/84 IDCR minke whale assessment cruises: estimating the hazard rate and effective strip width. Rep. int. Whal. Commn 35:315-18.
Hiby, A.R. and Ward, A.J. 1986a. Simulation trials of a cue counting technique for censusing whale populations. Rep. int. Whal. Commn 36:471-2.
Hiby, A.R. and Ward, A.J. 1986b. Analysis of cue counting and blow rate estimation experiments carried out during the 1984/85 IDCR minke whale assessment cruise. Rep. int. Whal. Commn 36:473-6.
Hiby, A.R. and Ward, A.J. 1988. Estimation of minke whale abundance using cue counting results from Shonan Maru No. 2 in Area II, 1986/87. Paper SC/40/Mi4 presented to the IWC Scientific Committee, May 1988.
Hiby, A.R., Ward, A.J. and Lovell, P. 1989. Analysis of the 1987 North Atlantic sightings survey: aerial survey results. Rep. int. Whal. Commn 39: 447-55.
Holt, R.S. 1984. Testing the validity of line transect theory to estimate density of dolphin schools. U.S. NMFS SWFC Admin. Rep. No. LJ-84-31, 56pp.
Holt, R.S. and Powers, J.E. 1982. Abundance estimation of dolphin stocks involved in the eastern tropical Pacific yellowfin tuna fishery determined from aerial and ship surverys to 1979. U.S. Dept of Commerce, NOAA Technical Memorandum No. NOAA-TM-NMFC-SWFL-23, 95pp.
Horwood, J. 1981. Results from the IWC/IDCR Minke Marking and Sightings Cruise 1979/80. Rep. int. Whal. Commn 31:287-314.
International Whaling Commission. 1982. Report of the Workshop on the Design of Sightings Surveys. Rep. int. Whal. Commn 32:533-49.
International Whaling Commission. 1984. Report of the Scientific Committee, Annex E1. Rep. int. Whal. Commn 34:77-101.
International Whaling Commission. 1986a. Report of the Workshop on Minke Whale Sightings. Rep. int. Whal. Commn 36:459-70.
International Whaling Commission. 1986b. Report of the Scientific Committee, Annex E. Report of the sub-committee on Southern Hemisphere Minke Whales. Rep. int. Whal. Commn 36:66-78.
International Whaling Commission. 1987. Report of the Scientific Committee, Item 6.2.3.(c) Rep. int. Whal. Commn 37:33.
Joyce, G.G. 1982. Blow patterns on sighting cues for censusing minke whales in Antarctic waters. Rep. int. Whal. Commn 32:787-90.
Joyce, G.G., Nakanishi, S., Hata, T., and Pastene, L. 1985. Preliminary evaluation of an apparatus for improved sighting angle evaluation. Rep. int. Whal. Commn 35:437-40.
Kasamatsu, F. and Kishino, H. 1986. Distribution of efforts in observer searching direction obtained in the independent observer experiment during the 1985/86 IWC/IDCR Southern Hemisphere minke whale assessment cruise. Paper SC/38/Mi9 presented to the IWC Scientific Committee, May 1986 (unpublished).
Kishino, H. 1986. On parallel ship experiments and the line transect method. Rep. int. Whal. Commn 36:491-5.

Kishino, H. and Kasamatsu, F. 1987. Comparison of closing and passing mode procedures used in sighting surveys. Rep. int. Whal. Commn 37: 253-8.
Ko, D., Zeh, J.E., Clark, C.W., Ellison, W.T., Krogman, B.D. and Sonntag, R.M. 1986. Utilisation of acoustic location data in determining a minimum number of spring migrating bowhead whales, unaccounted for by the icebased visual census. Rep. int. Whal. Commn 36:325-38.
Koopman, B.O. 1980. Search and Screening. Pergammon Press. 368pp.
Kraus, S.D., Gilbert, J.R. and Prescott, J.H. 1984. A comparison of aerial, shipboard and land-based survey methodology for the harbor porpoise, Phocaena phocaena. Fish. Bull. 81(4):910-3.
Leatherwood, S. 1982. Size of Bottlenose dolphin populations in India River, Florida. Rep. int. Whal. Commn 32:567-8.
Leatherwood, S., Awbrey, F.T. and Thomas, J.A. 1982. Minke whale response to a transiting survey vessel. Rep. int. Whal. Commn 32:795-802.
Leatherwood, S., Goodrich, K., Kinter, A.L. and Truppo, R.M. 1982. Respiration patterns and sightability of whales. Rep. int. Whal. Commn 32:601-13.
Leatherwood, S. and Reeves, R.R. 1982. Effects of varying altitude on aerial surveys of bottlenose dolphins. Rep. int. Whal. Commn 32:569-75.
Ljungblad, D.K., Moore, S.E. and Clarke, J.T. 1986. Assessment of bowhead whale feeding patterns in the Alaskan Beaufort and northeastern Chukchi Seas via aerial surveys, fall 1979-84. Rep. int. Whal. Commn 36:265-72.
Moore, S.E., Clarke, T. and Ljungblad, D.K. 1986. A comparison of gray whale (Eschrichtus robustus) and bowhead whale (Balaena mysticetus) distribution, abundance, habitat preference and behavior in the northeastern Chukchi Sea 1982-84. Rep. int. Whal. Commn 36:273-9.
Norris, K.S. and Harvey, G.W. 1979. A theory for the function of the spermeceti organ of the sperm whale. In: S.R. Geller, K. Schmidt-Koerig, G.J. Jacobs and R.E. Bellerville (eds.) Animal Orientation and Navigation. NASA Special Publication, 262pp.
Payne, P.M., Nicholas, J.R., O'Brien, L. and Powers, K.D. 1986. The distribution of the humpback whale Megaptera novaeangliae on Georges Banks in the Gulf of Maine, USA in relation to densities of the sandeel Ammodytes americanus. Fish. Bull. 84(2):271-8.
Quinn, T.J. 1979. The effects of school structure on line transect estimators of abundance. Contemporary Quantitative Ecology and Related Ecometrics. Statistical Ecology Series 12:473-91.
Quinn, T.J. 1985. Line transect estimators for schooling populations. Fish. Res. (Aust.) 3(3):183-200.
Raftery, A.E., Turet, P. and Zeh, J.E. 1988. A parametric empirical Bayes approach to interval estimation of bowhead whale, Balaena mysticetus, population size. Rep. int. Whal. Commn 38: 377-88.
Ramsey, F.L., Wildman, V. and Engbring, J. 1987. Covariate adjustments to effective area in variable-area wildlife surveys. Biometrics 43: 1-11.
Reilly, S.B., Rice, D.W. and Wolman, A.A. 1983. Population assessment of the gray whale Eschrichtius robustus from Caliornia USA shore censuses. 1967-1980. U.S. Natl. Mar. Fish. Serv. Fish. Bull 81(2):267-82.
Rugh, D.J. 1984. Census of gray whales at Unimak Pass, Alaska, November - December 1977-1979. pp. 225-48. In: M.L. Jones, S.L. Swartz and S. Leatherwood (eds). The Gray Whale. Academic Press, N.Y.
Schweder, T. 1977. Point process models for line transect experiments. pp.221-42. In: J.R. Barba, F. Brodeau, G. Romier and B. Van Cutsem (eds). Recent Developments in Statistics. North-Holland Publ. Co., New York, N.Y.
Scott, G.P and Gilbert, J.R. 1982. Problems and progress in the U.S. BLM-sponsored CETAP surveys. Rep. int. Whal. Commn 32:587-99.
Scott, M.D., Perryman, W.L. and Clark, W.G. 1985. The use of aerial photographs for estimating school sizes of cetaceans. Bull. Inter-Amer. Trop. Tuna Comm. 18(5):383-419.
Seber, G.A.F. 1982. The Estimation of Animal Abundance. Hafner Publ. Co., Inc., New York, NY. 654pp.
Skellam, J.G. 1958. The mathematical foundations underlying the use of line transects in animal ecology. Biometrics 14(3):385-400.
Smith, G.E.J. 1979. Some aspects of line transect sampling when the target population moves. Biometrics 35:323-20.
Smith, R.C., Dustan, P., Au, D., Baker, K.S. and Dunlap, E.A. 1986. Distribution of cetaceans and sea surface chlorophyll concentrations in the California USA current. Mar. Biol. (Berl.) 91(3):385-402.

Smith, T.D. 1981. Line transect techniques for estimating density of porpoise schools. J. Wildl. Mgmt 45:650-7
Smith, T.G., Hamill, M.O., Burrage, D.J. and Sleno, G.A. 1985. Distribution and abundance of belugas, Delphinapterus leucas and narwhals Monodon monoceros, in the Canadian High Arctic. Can. J. Fish. Aq. Sci. 42(4):676-84.

Sonntag, R.M., Ellison, W.T., Clark, C.W., Corbit, D.R. and Krogman, B.D. 1986. A description of a tracking algorithm and its application to bowhead whale acoustic location data collected during the spring migration off Point Barrow, Alaska. Rep. int. Whal. Commn 36:299-310.
Sonntag, R.M., Ellison, W.T. and Corbit, D.R. 1988. Parametric sensitivity of a tracking algorithm as applied to the migration of bowhead whales, Balaena mysticetus, near Pt. Barrow, Alaska. Rep. int. Whal. Commn 38: 337-47.
Swartz, S.L., Jones, M.L., Goodyear, J., Withrow, D.E. and Miller, R.V. 1987. Radio-telemetric studies of gray whale migration along the California coast: a preliminary comparison of day and night migration rates. Rep. int. Whal. Commn 37: 295-9.
Thomas, J.A., Fisher, S.R., Ferm, L.M. and Holt, R.S. 1986. Acoustic detection of cetaceans using a towed array of hydrophones. Rep. int. Whal. Commn (special issue 8): 139-48.
Thompson, D. and Hiby, A.R. 1985. The use of scale binoculars for distance estimation and a time lapse camera for angle estimation during the 1983/84 IDCR minke whale assessment cruise. Rep. int. Whal. Commn 35:309-14.
Turnock, D.J. 1987. The effect of responsive movement on abundance estimation using line transect sampling with application to Dall's porpoise. Unpubl. M.S. thesis, University of Washington, Seattle, WA. 115pp.
Ward, A.J. 1988. Accounting for the effect of pod size in the estimation of blow rates: analysis of data from the 1984/85, 85/86 and 86/87 IDCR minke whale assessment cruises. Rep. int. Whal. Commn 38: 265-8.
Ward, A.J. and Hiby, A.R. 1987. Analysis of cue-counting and blow rate estimation experiments carried out during the 1985/86 IDCR minke whale assessment cruise. Rep. int. Whal. Commn 37: 259-62.
Ward, A.J., Hiby, A.R. and Thompson, D. 1986. Photographic estimation of angles to sightings and distribution of sighting effort with angle - results obtained during the 1984/85 IWC/IDCR minke whale assessment cruise. Rep. int. Whal. Commn 36:477-9.

Watkins, W.A., Moore, K.E., Wartzok, D. and Johnson, J.H. 1981. Radio tracking of finback (Balaenoptera physalus) and humpback (Megaptera novaeangliae) whales in Prince William Sound, Alaska. Deep Sea Research 28:577-88.
Whitehead, H. 1982. Populations of humpback whales in the northwest Atlantic. Rep. int. Whal. Commn 32:345-53.
Whitehead, H. 1985. General behaviour of sperm whales off the Galapagos Islands, Feb-April 1985. Unpublished report.
Whitehead, H. and Gordon, J. 1986. Methods of obtaining data for assessing and modelling sperm whale populations which do not depend on catches. Rep. int. Whal. Commn (special issue 8): 149-65.
Winn, H.E. and Martin, A.M.H. 1983. Analysis of respiration rates of baleen whales off the northeastern U.S. coast. Abstract to 5th Biennial Conference on the Biology of Marine Mammals, Boston, Nov. 1983.
Withrow, D., Rice, D. and Wolman, A. 1983. Abundance and distribution of gray whales on the winter grounds and a comparison of the accuracy of aerial vs census techniques. Abstract to 5th Biennial Conference on the Biology of Marine Mammals, Boston, No. 1983.
Würsig, B., Dorsey, E.M., Fraker, M.A., Payne, R.S., Richardson, W.J. and Wells, R.S. 1984. Behaviour of bowhead whales, Balaena mysticetus, summering in the Beaufort Sea: surfacing, respiration and dive characteristics. Can. J. Zool. 62:1,910-21.
Würsig, B., Wells, R.S. and Croll, D.A. 1986. Behaviour of gray whales summering near St Lawrence Island, Bering Sea. Can. J. Zool. 64:611-21.
Yapp, W.B. 1956. The theory of line transects. Bird Study 3(2):93-104.
Zeh, J.E., Ko, D., Krogman, B.D. and Sonntag, R.M. 1986a. Statistical considerations in estimating the number of bowhead whales, Balaena mysticetus, from ice-based visual census data. Rep. int. Whal. Commn 36:317-23.
Zeh, J.E., Ko, D., Krogman, B.D. and Sonntag, R. 1986b. A multinomial model for estimating the size of a whale Balaena mysticetus population from incomplete census data. Biometrics 42(1):1-14.
Zeh, J.E., Turet, P., Gentleman, R. and Raftery, A.E. 1988. Population size estimation for the bowhead whale, Balaena mysticetus, based on 1985 visual and acoustic data. Rep. int. Whal. Commn 38: 349-64.

# Molecular Techniques for Examining Genetic Variation and Stock Identity in Cetacean Species 

A. Rus Hoelzel and Gabriel A. Dover<br>Department of Genetics, University of Cambridge, Downing Street, Cambridge CB2 3EH, England

Accepted May 1988


#### Abstract

This report surveys currently available molecular methods for the measurement and analysis of naturally occurring genetic variation, with particular reference to whales and other mammals. We concentrate on three cellular components that are amenable to molecular analysis, namely proteins, mitochondrial DNA and nuclear DNA. Sections 2, 3 and 4 describe the nature of variation in these three components respectively, with emphasis on the utility of such variation for the identification of taxa at increasing taxonomic distances (individuals, sibships, pedigrees, breeding groups, populations, species and genera). At the same time methods are described which assess the genetic distances between taxa, their times of separation and the breeding and migratory habits of the animals under study. The precise genetic identification of individuals, revealed by new methods of DNA 'fingerprinting' of several distinct nuclear DNA components, could yield detailed information on family sizes, the reproductive success of individuals and population dynamics in general.

The report is aimed to be self-contained in giving users a broad outline of the forces known to shape the genetic architecture of populations, together with details of the analytical procedures available to understand the raw data. Some assessment of the strengths and weaknesses of each procedure is given in appropriate sections. Section 1 provides a general survey of the nature of mutations, and the modes and rates of their dissemination through a population. Recent findings on the unexpected non-Mendelian behaviour of DNA lengthens the list of types of mutation and the ways in which they may accumulate. These new genetic processes and their influence on population differentiation are described in detail in Sections 1 and 4. The more traditional approaches to the understanding of natural variation, on which there is a considerable body of literature, are detailed in Sections 2 and 3. Section 5 outlines some of the practical methods for collecting and preserving samples, and for uncovering natural variation at the molecular level amongst whales.

From this survey we conclude that breeding behaviour, population dynamics, census and stock identification are feasible through the molecular analysis of hypervariability ('fingerprinting') in the mitochondrial genome, the nuclear ribosomal RNA multigene family and several nuclear minisatellite DNA families, in preference to the analysis of protein variation, gross DNA divergence, DNA point-mutations (restriction fragment length polymorphisms, RFLPs) and satellite DNA families. Small skin samples collected from live, unrestrained animals are sufficient for all applications involving population surveys.


## 1. GENERAL OUTLINE OF GENETIC VARIATION

## I. 1 Introduction

The need for a review of the use of biochemical techniques for determining stock identity was identified by the IWC's Scientific Committee at its meeting in June 1986. This developed from discussions of the implications for management of discrepancies between the boundaries of 'biological' and 'management' stocks at the Special Meeting of the Scientific Committee on Planning for a Comprehensive Assessment of Whale Stocks (SC/38/Rep 1). In particular, the Scientific Committee recognized that it was necessary to know what information could be obtained about the genetic nature of stocks before the simulation studies recommended in Annex D of SC/38/Rep 1 were conducted. Our report provides a detailed survey of techniques from molecular biology which are suitable for investigating and analysing the genetics of cetacean stocks.
The Scientific Committee has distinguished between genetic (or biological) stocks, dynamic stocks, and management units. We assume that a dynamic stock is the fundamental unit described by a population model or assessment procedure. A management unit is the group of whales occurring within a specific geographical boundary which is actively or potentially exploited by the member nations of the IWC. These are the individual stocks whose status is assessed by the Scientific Committee. A genetic stock is a genetically differentiated population within a species. We will be more explicit about how such a population can be identified later.

The central question in Annex D of SC/38/Rep 1 is: to what extent do genetic stocks and management units
coincide? This is really two questions. (1) Are the whales found in two geographically distinct management units from two different genetic stocks? If the whales in the two units cannot be distinguished genetically this does not necessarily mean that they are all part of the same genetic stock. The next question is therefore: (2) Are individuals from more than one genetic stock present in a particular management unit? If the answer to either question is 'yes' it is then useful to ask: what level of interchange may have occurred between the different genetic stocks? For some cetacean species the location of breeding aggregations is known. In these cases it may be useful to determine whether the different aggregations are also discrete genetic stocks. This is, of course, a special case of question 1.

The report that follows is a detailed description of the strengths and weaknesses of molecular methods for uncovering, analysing and interpreting genetic variation. As such the report uses genetical and statistical terms which are unavoidable but which are familiar to the specialist geneticist who might wish to use this report as a basis for an experimental investigation into the genetic basis of stocks and populations of cetacean species. The report is, however, also directed at the general marine mammalogist and we have tried to make it as comprehensible as possible. As a guide to understanding the report we have written two appendices. Appendix 1 is a short summary of the report; a self-contained description of our recommendations. Appendix 2 is a glossary of words with short definitions. Many of the more unfamiliar molecular terms are also described in full in the text.

The conservation and management of cetaceans depends on an effective method for the identification and
description of breeding populations. Genetic variation within a species is discontinuous, and removal by overexploitation or habitat disruption of one regional population may eliminate variation important to the species as a whole.

Cetaceans are highly mobile and live in a relatively homogeneous environment. Further, for most species very little is known about patterns of distribution and breeding behaviour. Taken together these factors make speculation about the delineation of breeding populations difficult. Species that have been harvested by the whaling industry have been studied in more detail. However, information from catch records and mark-recapture census techniques do not provide unequivocal data on population boundaries or dispersal. The greatest potential for uncovering these parameters derives from molecular techniques, especially in combination with observational data.

Some cetacean populations have been reduced to low levels (e.g. right whales, see IWC, 1986). This raises the concern that inbreeding depression: a reduction in fitness through the exposure of deleterious homozygous recessive genes, could further reduce numbers to a level where recovery is not possible. Various factors are important to determine whether a species is at risk. For example, the proportion of breeding age animals successfully breeding in a given season (effective reproductive population size), dispersal frequency and range, and levels of heterozygosity. Modern molecular techniques allow a detailed analysis of breeding behaviour and differential breeding success, as well as improving estimates of stock boundaries and dispersal patterns, which facilitate the assessment of this problem in small populations.

Genetic variation can be measured at all levels, from nucleotide sequence to the phenotypically expressed characters that distinguish larger taxonomic units. Each level requires its own method of analysis, although they overlap. Within and between population variation can be measured in the structure of DNA, proteins, chromosomes, morphology and behaviour. Chromosomal structure has been shown to be a useful tool for describing cetacean systematics (see Kulu, 1972; Arnason, 1974). This review will emphasize the use of molecular characters such as DNA and protein variability.

### 1.2 Nature of mutations

Genetic variation accumulates primarily by mutation at the DNA level. The DNA molecule, which consists of a long chain of paired nucleic acid bases, can be altered in a variety of ways. The best understood, but by no means the most prevalent, involves the substitution of one base for another (point mutation). Point mutations accumulate primarily by copy error (during replication) and by induced mutation (from 'mutagens' in the environment).

Other types of mutation involve the removal or addition of stretches of DNA (which may vary in length from two to several thousand nucleotides); the inversion of a length of DNA at a given position; and the movement of lengths of DNA from one position to another. Such mutations are caused by mechanisms of DNA turnover known as slippage, unequal crossing over, gene conversion, and transposition, amongst others. Their modes of operation are described in Section 1.3. Collectively they place the nuclear (and sometimes mitochondrial) genome in a state of flux, generating a wealth of genetic variability. Such mechanisms are also involved in the dissemination of
point-mutational and molecular rearrangements through a population, with the passing of the generations (see Section 1.3).

Variation in a population also arises as a consequence of the sexual process. This is due to the independent segregation of non-homologous chromosomes, and crossing over between homologous chromosomes. The effect of these processes on variation has been reviewed by Lewontin (1974). It has long been considered that the creation of new combinations of alleles by recombination would facilitate an organism's capacity to adapt to changing environments (e.g. Carson, 1959). However, advantageous genotypes will break apart as readily as they form (Eshel and Feldman, 1970).

### 1.3 Dissemination of mutations in populations

Essentially, there are three ways by which mutations can spread in a sexual population. These are natural selection, genetic drift and molecular drive (Dover, 1982; 1986a). Natural selection is a consequence of differences in the extent to which genetically distinct individuals interact with their environment: an interaction affecting their relative reproductive success. Genetic drift is a consequence of sampling error of genes due to the continual stochastic gain and loss of gametes and individuals in a population.

Molecular drive (Dover, 1982) is a term used to describe the various mechanisms of DNA turnover that operate independently of selection and drift. We will describe these in some detail as they are involved in the generation of variation in genomic regions most useful to the assessment of kinship and genetic distance between populations. Each mechanism of turnover is given a different name reflecting its mode of operation. Briefly these are as follows:

## (i) Transposition

This is the most commonly understood mechanism. It involves the movement of a length of DNA from one position to another in the genome. There are two types of transposition: the first involves the duplication of a length of DNA followed by its insertion elsewhere (duplicative transposition); the second simply involves excision and reinsertion, (non-duplicative transposition). These mobile genetic elements have a variety of structures indicating that at the molecular level there are a variety of means by which DNA can move around the genome. All new insertions, however, cause a small duplication of the sequences at the site of insertion. Hence, mobile elements can often be recognised by the presence of two direct short repeats (usually 5-10 bases long) flanking the element, with only one copy in the genome before insertion.

Variation induced by mobile elements can take several forms. If the element moves into a gene or its flanking controlling sequences then a mutation can ensue. Further, it is known that excision of an element is imprecise in that a few extra bases may be left behind as a 'footprint' at the site of excision. Such 'footprints' can affect the expression of a gene and lead to the gradual accumulation of a multiple allelic series at a locus (Coen, Carpenter and Martin, 1986). If an element moves into non-genic sequences (of which there are many in most genomes - see below), then the distances between restriction sites are altered. This is detectable by standard molecular methods (see Section 3.1). In general, the rates of transposition are low (from $10^{-2}$ to $10^{-4}$ events per generation). Hence, different populations can be identified by different but relatively invariant positions of a family of mobile elements.

Mobile elements represent a clear case where the segregation behaviour of the DNA does not follow that of the chromosomes. Mendel's Laws are derived from the random assortment of chromosomes at meiosis and the random fusion of haploid chromosome sets at fertilisation. The degree to which the DNA and chromosomes are out of synchrony varies from one family of mobile elements to the next; nevertheless the ability of elements to make extra copies of themselves and move around the genome ensures the accumulation of an element in a population. As with all other DNA turnover mechanisms (see below) such accumulation is operationally distinct from the accumulation of variants under natural selection or genetic drift, based as it is on internal mechanisms of gain and loss of genetic material. Detailed examination of genetic variation in some multigene families (a set of genes descended by duplication and variation from some ancestral gene) reveals the extent to which this process interacts at a higher level with natural selection (Dover and Flavell, 1984; Tautz et al., 1987; Arnheim, 1983; Gerbi, 1985); and see later.

## (ii) Unequal crossing over

A crossover can occur between two chromatids or between two chromosomes (homologous and sometimes non-homologous), when there is not complete alignment between the two structures. After crossing over, one structure gains extra genetic material and the other suffers a corresponding loss [see Fig. 2, and Section 4.2(i)]. Unequal crossing over is a means of generating duplicate genes, as seen for example in some of the mammalian globin genes. The continual activity of unequal crossing over can produce further rounds of duplication leading to long tandem arrays of a given gene; and hence generating a multigene family. There are many gene families that have been generated in this way, the most notable being the genes for the ribosomal RNAs (18S, 28S and 5S RNAs, see 4.2); the five histone genes (often in repetitive units each of which contains all five genes); and the several 'variable' and 'constant' genes of the mammalian immune superfamily of genes (for reviews see Ohta, 1980; 1983; Dover, 1982; 1986a; Arnheim, 1983; Long and Dawid, 1980; Hood, Kronenberg and Hunkapiller, 1985; Kedes, 1979; Hood, Campbell and Elgin, 1975).

Once a gene family is established, unequal crossing over is involved with its maintenance, in the sense that the mechanism ensures that genetic variation between member genes is continually reduced, (Smith, 1974; Ohta, 1980; Dover, 1982). For example, if a mutation occurs in one of a hundred genes in an array, then there is some probability that after many rounds of unequal crossing over, that the mutant gene will replace the original array. This is because unequal crossing over induces a continual process of stochastic gain and loss. In the early stages, when the mutant gene is rare, there is a high probability that it will be lost. However, should it begin to diffuse through the array, then the probability of it continuing to do so increases. The process of spreading a mutant gene through a gene family is exactly analogous to the process of diffusion of a neutral allele by genetic drift through a population. The latter process is consequential on the stochastic gain and loss of gametes and individuals.

Unequal crossing over between chromosomes ensures that a mutant copy spreads to all chromosomes on which the family of genes is situated. All chromosomes then enter
different individuals after the sexual process, in each of which the homogenisation process can continue. Hence, a combination of unequal crossing over and the sexual process ensures that all arrays of a gene in a population are being similarly homogenised. Eventually all member genes of a given family acquire high levels of genetic identity by unequal crossing over, which is one of the several turnover mechanisms that underpin molecular drive.

Evidence that this has happened in most coding and non-coding DNA families derives from a common observation that a given family in a species contains species-diagnostic mutations, which have been homogenised throughout the family, (or some subsection of it). Detailed studies of large DNA families reveal all the stages of transition during this process (Strachan et al., 1982; Strachan, Webb and Dover, 1985). The utility of such homogeneity patterns (called concerted evolution) for the study of natural variation is described in Sections 4.1 and 4.2.

Unequal crossing over (in conjunction with slippage see below) is responsible for the observed hypervariability in some DNA families such as the ribosomal DNA (rDNA) and the minisatellites used for DNA 'fingerprinting'. This variability is due to the differences in the number of copies of a tandem array of genes, as a consequence of continual gain and loss by unequal crossing over or slippage (see Fig. 2, Section 4.2). Copy-number variability should not be confused with the low variability in sequence between copies due to the homogenising consequences of such turnover mechanisms. These aspects are described in detail in Sections 4.2 and 4.3.

## (iii) Slippage

The precise molecular events of slippage (sometimes known as slippage-replication) are not known. It is recognised, however, that many regions of nuclear genomes are composed of very short (on average less than 10 base pairs) motifs of DNA that are in tandem arrays (pure simplicity) or scrambled one with another (cryptic simplicity), (for review see Tautz, Trick and Dover, 1986). The numbers of copies of any given motif in any defined region are much higher than would be expected to occur by chance in a random sequence of the same length and composition of As, Ts, Cs and Gs. Both pure and cryptic simplicity are considered to be due to the propensity for the two strands of the DNA helix to slip against each other, creating a gap on one side and a buckle (loop) on the other. Repair of such lesions or structures can lead to a gain or loss of short motifs of DNA. Detailed computer analyses of many genes, when compared between species, indicate that slippage-generated variation is widespread in exons, introns and flanking sequences and that it is being produced at a faster rate than the point mutation rate. Slippage-like mechanisms of turnover are being recognised as a major source of genetic variation (Tautz et al., 1986; Bird, 1986). It is not yet clear whether the hypervariable minisatellites in the human genome are generated by slippage or unequal crossing over. Operationally the two processes are very similar, with the latter generally considered to be unable to cope with the very short repetitive motifs comprising the simple sequence regions of DNA.
It is assumed that slippage, unlike unequal crossing over, occurs only within a chromatid (double helix). As such it can be involved only with the comings and goings of motifs
in a single replicating lineage, and cannot lead to the spread of any given motif to other lineages. Recent studies on slippage-generated repetition within the silk moth chorion (egg-shell) genes show that another mechanism, gene conversion (see below) is responsible for spreading the motifs to other members of the family of chorion genes no matter what their chromosomal location might be (Eickbush and Burke, 1986; Dover, 1986b). Another recent study on the rDNA in species of Drosophila (Tautz et al., 1987) shows that turnover by slippage is operating within each of a small array of subrepeats that lies within the longer rDNA repeating unit (see Fig. 1, Section 4.2). The array of subrepeats (each 90 base pair long) is being homogenised by unequal crossing over in some species. However, in other species the much faster rate of slippage operating on smaller units of DNA is destroying the once homogeneous array of subrepeats. Such studies in silk moths and Drosophila are two examples out of many which illustrate the simultaneous operation of different turnover mechanisms within the same region of DNA. Complex patterns of DNA divergence emerge when such mechanisms operate on different unit lengths of DNA, and at different rates (Dover, 1987).

## (iv) Gene conversion

Analysis of non-Mendelian patterns of gene segregation in meiotic tetrads of fungal species reveals the phenomenon of gene conversion (Whitehouse, 1983). This is a mechanism which involves the non-reciprocal transfer of sequence between copies (alleles or non-alleles) of a gene. That is, starting with two slightly different copies, gene conversion leads to two identical copies. This process is thought to be due to the invasion of a double helix of one member gene by a single strand of the helix of another. After much twisting and turning the resultant heteroduplex is repaired to give rise to a stably base-paired homoduplex. The direction of repair can be arbitrary (unbiased gene conversion) in which case a heteroduplex of composition $A a$ can be repaired to either $A A$ or $a a$. If the repair is more frequently in the direction of either $A$ or $a$ then it is said to be biased.

Gene conversion can involve regions of DNA from just a few bases to tens of thousands of bases. Although by definition it is a process of homogenisation, nevertheless it can lead to genetic variation if the unit of DNA under comparison between taxa is longer than the gene conversion domains within it. In such an instance the unit of DNA becomes a mosaic of different conversion domains, and each unit in the separate taxa is a differently composed mosaic. Much of the high variability in the several genes involved with the mammalian immune system has arisen by such a disparity between the length of the gene and the conversion domains, (for reviews see Baltimore, 1981; Dover and Strachan, 1987).

Gene conversion can be involved with the spread of mutations through gene families and through sexual populations, (Ohta, 1980; Dover, 1982; Lamb and Helmi, 1982; Nagylaki and Petes, 1982), and hence affects the expected levels of population heterozygosity. This is true for both single-copy and multiple-copy genes. In fact, given the widespread observation of slippage operating in some regions of genes, coupled to gene conversion, it is unlikely that true Mendelian genes exist which do not contain any internal repetition and whose mutant alleles rely solely on selection or drift for their increased representation in a population. The analysis of sequences
of all available gamma-globin alleles in humans indicates that gene conversion is operating seven to ten times faster than the point-mutation rate (Smithies and Powers, 1986).

A bias in gene conversion can initially accelerate the rate of divergence of a given gene between taxa, if the bias is for different gene variants in the two taxa. If, however, in each of two further taxa derived from one of the original taxa, the same bias were to be maintained, then there would be a retardation in the rate of divergence between the two new taxa. Hence, both conservation and divergence of sequence can result from a gene conversion bias depending on its mode of operation and the time of origin of the taxa under review (Dover, 1987). Conservation and divergence of sequences do not necessarily imply, in the absence of the relevant evidence, the operation of selection or drift respectively.

## (v) RNA-mediated transfers of genetic information

Some proportion of the available genetic variation in nuclear genomes is due to the turnover of DNA sequences via their RNA intermediates. This is a consequence of the presence of reverse transcriptase which transcribes RNA into its complementary DNA (cDNA), followed by the reinsertion of the cDNA into the genome at many different loci. Many processed pseudogenes (i.e. genes without introns and the $5^{\prime}$ and $3^{\prime}$ control sequences) arise in this way. Some very large DNA families such as the 500,000 copies per individual human of the 'Alu' family, in addition to many other repetitive families, arise via their RNA intermediates. Such a mechanism can lead to a relatively rapid accumulation of repetitive elements and pseudogenes because of the vast numbers of RNAs per nucleus, leading to large differences in the copy-number of a given repetitive family, even between closely-related species.

Molecular drive, selection and drift are all operationally distinct but superimposed one upon another, leading to complex patterns of population change and differentiation. Unravelling all three processes becomes a necessity when trying to assess the nature and significance of genetic variation, in particular at the DNA level. Fortunately, the difficulties in quantifying the relative contributions of the three processes to observed diversity in any given genomic component does not hamper the exploitation of this diversity for the correct identification of hierarchical levels of taxa (see Section 4).

### 1.4 Mutation rates

In general, the lowest rates of nucleotide substitution occur in coding sequences. However, within this category rates are highly variable. They range from $0.004 \times 10^{-9}$ substitutions per year in histone IV (Dayhoff, 1972; Wilson, Carlson and White, 1977) to $2.8 \times 10^{-9}$ in interferon A (see Li, Lou and Wu, 1985a for a review on the range of non-synonymous substitution rates in the first and second positions of codons). The average rate for mammals is $0.88 \times 10^{-9}$ substitutions per non-synonymous site per year (Li et al., 1985a). The substitution rate at synonymous sites (especially at the third codon position, which can vary without leading to an amino-acid change) has been suggested to be quite uniform (Miyata et al., 1980; Hayashida and Miyata, 1983). This is about $5.5 \times 10^{-9}$ substitutions per year. However, Li et al. (1985a) suggest greater variation (from 1.7 to $11.8 \times 10^{-9}$ substitutions per year). In either case, the average: about $5 \times 10^{-9}$ substitutions per year, is five times higher than average
non-synonymous substitution rates. The effect of inconsistent substitution rates on models for quantifying population diversity are discussed below.

Unequal crossing over is one of two ways for generating an extra copy of a gene (gene duplication). If either copy acquires a nonsense codon (caused by a misplaced end-signal codon) or frameshift mutation then the gene produces a non-functional product. Such sequences are referred to as pseudogenes. Because there should be no functional constraints on pseudogenes, it has been postulated that they will evolve at a rate equivalent to the mutation rate (e.g. Li, Wu and Lou, 1985b). This is on the assumption that pseudogenes are independent and do not engage in any of the DNA turnover mechanisms referred to above. Nucleotide substitution rates in pseudogenes were found to be about the same as the rate for synonymous substitutions (average $=4.85 \times 10^{-9} ;$ Li et al. , 1981; Li, 1983; Gojobori and Nei, 1984; Li et al., 1985b).

Eukaryotic genes in nuclear genomes are split into coding segments (exons) separated by stretches of non-coding DNA (introns) which are transcribed into RNA, but are not translated into a polypeptide chain. Further, genes begin and end with non-transcribed flanking sequences. Li et al. (1985b) have compared substitution rates at these various sites and found intron regions to evolve at $76 \%$ of the rate for pseudogenes, and flanking regions at $50-90 \%$. This is considerably higher than the average rate for coding regions, which evolve at $18 \%$ the rate of pseudogenes.

As described in Section 1.3, many genes in eukaryotic nuclear genomes exist in multiple copies (multigene families) due to repeated duplications by unequal crossing over and similar amplification mechanisms. The rate of divergence in multigene families (such as those coding for histones, immunoglobulins and ribosomal RNAs) is complicated by the fact that a mutation occurring in one member gene can spread to other copies of the family, by any one of several DNA turnover mechanisms. The family or some subsection of it becomes homogenised and evolves as a unit (Arnheim, 1983; Dover, 1982). Furthermore, multiple genes that are in a tandem array are often separated by spacers which experience weaker functional constraints than the genes themselves. No strict comparisons can be made, therefore, between rates of divergence in so-called 'single-copy' genes (i.e. two per diploid individual) and multigene families. With this proviso, Ohta (1980) has compared rates of divergence within the immunoglobulin multigene family between humans and rabbits. She describes a rate of $0.7 \times 10^{-9}$ substitutions per nucleotide site per year in the coding regions and $1.8 \times 10^{-9}$ for the spacer regions (which corresponds to the high end of the spectrum for structural genes). This higher rate of divergence might be due to the homogenisation consequences of DNA turnover, which are known to be operating at rates from $10^{-2}$ to $10^{-4}$ per kilobase per generation, (Coen, Strachan and Dover, 1982a; Coen, Thoday and Dover, 1982b; Jeffreys, Brookfield and Semeonoff, 1985). These rates are at least two orders of magnitude faster than the mutation rate per kilobase per generation, (Dover, 1982; 1986a).

The highest known evolutionary rates are in the satellite DNA sequences such as the hypervariable minisatellite regions described by Jeffreys and co-workers (Jeffreys, Wilson and Thein, 1985b). Nucleotide substitution rates in these regions are over an order of magnitude higher than the average for coding regions (about $2.0 \times 10^{-7}$ ). The
assessed rate of unequal crossing over or slippage which is responsible for the evolution of human minisatellites (Section 4.3) is approximately $0.5-1.5 \times 10^{-4}$ per kilobase per gamete (Jeffreys et al., 1985). This disparity in rates is sufficient for new mutations to be homogenised amongst the repetitive units of the minisatellite DNA.

In summary, nucleotide substitution rates in nuclear DNA vary from about $10^{-12}$ (histone IV) to $10^{-7}$ (minisatellites). The observed level of variation in a population is the result of this gradual accumulation of mutations (countered by chance backward mutations), coupled to the actions of natural selection, genetic drift and molecular drive. Genetic drift is most likely to be a factor in finite populations (less than $10^{6}$ individuals), increasing in importance with decreasing population size. It is a stochastic effect which either brings new mutations to fixation (represented in all individuals) or eliminates them by chance.

### 1.5 Analysis of genetic variation

The analysis of variation at the molecular level began at the turn of the century with studies on blood types in humans (Landsteiner, 1900). However, the more extensive characterisation of variation in other species did not begin until 1966 with the application of gel electrophoresis (Lewontin and Hubby, 1966; Harris, 1966). In the first ten years after the technique was introduced, genetic variation at loci coding for proteins was described for nearly 250 species (where 14 loci or more were investigated; see reviews by Powell, 1975; Selander, 1976; Nevo, 1978). It became apparent that there is extensive genetic variation in natural populations. To date well over 1,000 species have been investigated (see Nevo, Beilles and Ben-Shlomo, 1983). The emphasis in marine species has been on invertebrates (e.g. Battaglia and Beardmore, 1977; Flowerdew, 1983), cod and salmonids (see review by Allendorf and Utter, 1979).

The species or population under study is usually described in terms of the proportion of polymorphic loci per population ( P ) and heterozygosity ( H : heterozygosity per locus per individual). At a gene locus with two alleles (variants, A and a; possible genotypes, AA, Aa and aa), assuming random mating, the allele frequency is defined below. If the total number of individuals in the population is N let the number possessing each genotype be $\mathrm{N}_{\mathrm{AA}}, \mathrm{N}_{\mathrm{Aa}}$ and $\mathrm{N}_{\mathrm{aa}}$. The frequency of the A allele in the population will be

$$
\mathrm{p}=2 \mathrm{~N}_{\mathrm{AA}}+\mathrm{N}_{\mathrm{Aa}} / 2 \mathrm{~N}
$$

and the frequency of the a allele will be

$$
\mathrm{q}=2 \mathrm{~N}_{\mathrm{aa}}+\mathrm{N}_{\mathrm{Aa}} / 2 \mathrm{~N}
$$

The heterozygosity $(\mathrm{H})$ is the proportion of Aa genotypes, which according to the Hardy-Weinberg law is 2pq, ( $p^{2}+2 p q+q^{2}=1$ ). In multiple allele systems heterozygosity can be defined as:

$$
1-\Sigma x_{i}{ }^{2}
$$

where $x_{i}$ is the frequency of the $i$ 'th allele at a given locus. ' H ' for a species is simply the average heterozygosity for all loci investigated. In non-random mating populations the above quantity is not related to the frequency of heterozygotes, but is nevertheless a good measure of genetic diversity (see Nei, 1975).

Nevo et al. (1983) computed average values for polymorphism ( P ) and heterozygosity (H) based on 968 plant and animal species: $P=0.284 \pm$ SD 0.197 and
$\mathrm{H}=0.073 \pm$ SD 0.076 . The average for 551 vertebrate species was lower: $\mathrm{P}=0.226 \pm$ SD 0.146 and $\mathrm{H}=0.054 \pm \mathrm{SD}$ 0.059 . Mammals were at the low end of that group: $\mathrm{P}=0.191 \pm \mathrm{SD} 0.137$ and $\mathrm{H}=0.041 \pm \mathrm{SD} 0.035$.

Caution is necessary in the interpretation of average P values however, as sample sizes vary greatly between studies as do criteria for polymorphism. A locus is most commonly defined as polymorphic if the most common allele frequency is 0.99 or less, but other criteria are used and not always stated in the published report. Further, when a small number of loci are investigated (e.g. 24 loci; see Nei, Maruyama and Chakraborty, 1975) the estimate of average heterozygosity is subject to a large standard error (Nei and Roychoudhury, 1974a). In the review published by Nevo (1978), 24 or less loci were investigated in $74 \%$ of the studies. Nei (1975) recommends that estimates of average heterozygosity be conducted on as many loci as possible, ideally a random sample of the genome. The number of individuals on the other hand, can be as low as 20 (Nei, 1978).

Selection will affect the expression of new variation through the differential survival of favourable and deleterious phenotypes. Selection can also maintain variation if the heterozygous condition is favoured at a given locus (Rendel, 1953: for a review see Falconer, 1981) or through various mechanisms of 'balanced' selection (see Hartl, 1980). However, selection can only act directly on DNA sequences that are expressed phenotypically.

The existence of DNA sequence variation can be a powerful tool in the characterisation of populations. A number of techniques have been developed to assess levels of genetic variation. The degree and type of variation depends very much on what part of the genome is being investigated. Examining variation in proteins (and therefore in the genes that code for them) has been by far the most common technique. The strength of this approach is its emphasis on DNA sequences that are expressed phenotypically. This allows an investigation into the role of selection in the evolution of gene loci. The procedure is also inexpensive and relatively easy to conduct (especially horizontal starch gel electrophoresis). However, protein studies show variation in the most conserved class of DNA (single copy coding sequences). Further, the highly variable segments within the structural gene (introns, flanking sequences and synonymous third codon sites) are not detected by this technique.

A number of new recombinant DNA techniques are now available for the analysis of more variable regions of the genome. Often these procedures involve the isolation of a DNA sequence which is then radioactively labelled and used to 'probe' the genome for similar sequences. This work usually requires greater expense and technical expertise, but a far higher level of resolution is gained, (see Sections 3 and 4).

## 2. PROTEIN VARIATION IN NATURAL POPULATIONS

### 2.1 Population genetic studies on Cetacea

Enzyme variation in marine mammals has been described for eight species of pinniped where the number of loci investigated was 15 or more (see Testa, 1986) and thirteen cetaceans (Simonsen, Born and Kistensen, 1982; Wada, 1983a; Wada, 1983b; Shimura and Numachi, 1987; Winans and Jones, in press; Duffield, unpublished), see Table 1. For the pinnipeds, mean $\mathrm{P}=0.076 \pm \mathrm{SD} 0.06, \mathrm{H}=0.019 \pm$

SD 0.01 . Heterozygosity ranged from a low of 0.000 for the northern elephant seal, Mirounga angustirostrus (Bonnell and Selander, 1974) to 0.033 for the walrus, Odobenus rosmarus (Simonsen, Kapel and Larsen, 1982b). The northern elephant seal study investigated 24 presumptive gene loci for 159 individuals from five rookeries.

Table 1
Enzyme variation in cetaceans. References: $\mathbf{A}=$ Simonsen et al., 1982a; B = Danielsdottir et al., 1988; C = Wada, 1983b; D = Shimura and Numachi, 1987

| Species | Number of loci investigated | N | P | $\mathrm{H} \pm$ SD | Ref |
| :---: | :---: | :---: | :---: | :---: | :---: |
| B. acutorostrata | 15 | 64 | 0.095 | $0.040 \pm 0.010$ | A |
| B. physalus | 26 | 47-218 | 0.212 | 0.055 | B |
| S. coeruleoalba | 15 | 40 | 0.130 | $0.021 \pm 0.008$ | C |
|  | 19 | 370 | 0.263 | $0.089 \pm 0.160$ | D |
| P. crassidens | 19 | 31 | 0.211 | $0.051 \pm 0.092$ | D |
| G. macrorhynchus | S 19 | 39 | 0.263 | $0.054 \pm 0.106$ | D |
| S. attenuata | 19 | 183 | 0.263 | $0.089 \pm 0.170$ | D |
| T. truncatus | 19 | 35 | 0.105 | $0.039 \pm 0.113$ | D |
| L. obliquidens | 19 | 30 | 0.316 | $0.093 \pm 0.182$ | D |
| S. bredanensis | 19 | 29 | 0.053 | $0.007 \pm 0.024$ | D |
| P. dalli | 19 | 483 | 0.421 | $0.154 \pm 0.184$ | D |
| B. bairdii | 18 | 12 | 0.056 | $0.016 \pm 0.069$ | D |
| P. electra | 19 | 6 | 0.105 | $0.035 \pm 0.108$ | D |
| P. phocoena | 18 | 3 | 0.167 | $0.047 \pm 0.111$ | D |

The cetacean species showed somewhat higher levels of enzyme variation, although the data are difficult to interpret as the number of loci investigated was fairly small. It has been demonstrated that heterozygosity estimates increase for the same population when fewer loci are studied (Nei, 1975). Forty striped dolphins (Stenella coeruleoalba) were investigated at 15 presumptive loci and found to have $\mathrm{P}=0.130$ ( $95 \%$ criterion for polymorphism) and $\mathrm{H}=0.021 \pm$ SE 0.008 (Wada, 1983b). These data are different from data of Shimura and Numachi, presented in Table 1 , possibly due to difference in sample size. Sixty four minke whale liver samples were examined for variation at 15 presumptive loci and a $\mathbf{P}=0.095$ was found ( $99 \%$ criterion for polymorphism), $\mathrm{H}=0.046 \pm$ SD 0.01 (Simonsen et al., 1982a). These levels of variability are lower than typically found in most other mammalian species. Consistently, Sharp (1975; 1976) reported low levels of blood protein polymorphisms in several races and species of delphinid cetaceans, as did Borisov (1981a and 1981b) for mysticetes. Shimura and Numachi (1987) surveyed 12 odontocete species at 19 loci and found higher levels of variation. Averaged over all species, the proportion of polymorphic loci was $0.164 \pm 0.112$ and heterozygosity was $0.047 \pm 0.035$. However, sample sizes varied from 3 individuals (Phocoena phocoena) to nearly 500 individuals (Phocoenoides dalli). If only the eight species represented by 20 or more individuals are considered, then the proportion of polymorphic loci is $0.237 \pm 0.108$ and heterozygosity is $0.072+/-0.042$, which is considerably higher than published averages for mammals (see above).

Wada and Numachi (1979) applied chi-squared analysis to an isozyme study of liver samples from nearly 2,500 Antarctic minke whales taken in the 1975-1977 whaling seasons. Samples were collected from $10^{\circ}$ square regions spread out over most of the Antarctic. Three polymorphic
enzymes (6-PGD, SDH and GOT) were used to compare samples from different regions. Sample sizes in some squares were quite small, so the $2110^{\circ}$ samples were clumped into five large samples roughly corresponding to existing Area boundaries (e.g. see IWC, 1987, p. 404). If this was not done, sample error was high and no significant differences could be detected. Wada and Numachi support this clustering by demonstrating that the chi-squared goodness of fit test on the Hardy-Weinberg equilibrium did not show a significant deviation for each of the five samples (see Section 2.8). Although Van Beek and Van Biezen (1982) suggested that the chi-squared test is not suitable for this comparison due to small numbers in many of the classes, the maximum likelihood test they recommended produced similar results.

Significant differences in allele frequencies at 6-PGD and SDH loci were described by Wada and Numachi (1979) from comparisons of samples collected on opposite sides of $130^{\circ}$ E in the Antarctic Ocean. However, Van Beek and de la Mare (1981) pointed out that a chi-squared test on a contingency table of alleles against 'Areas' for each enzyme was not significant at the 0.05 level in either case. Therefore the null hypothesis that the data are homogeneous is accepted. They suggest that this renders the pairwise comparisons by which Wada and Numachi found significant differences 'a doubtful procedure' (a similar view was expressed by Horwood, 1980). This point is illustrated by example for the 6-PGD data. Within homogeneous data sets the probability of making a type one error (the chance that a true hypothesis was rejected) is compounded by the number of comparisons. For the 6-PGD data there are 40 contrasts. The probability that there will be a type one error is one minus the probability of no type one error. If the error level per comparison is set at 0.05 , then the probability of a type one error overall is $1-(0.95)^{40}=0.87$. Further, they suggest that pair-wise comparisons of different alleles at the same locus cannot be counted as independent. From these problems they conclude that the data presented by Wada and Numachi 'cannot be considered to discriminate separate stocks of minke whales in the Antarctic'.

Wada (1982) re-analysed the data by the G-statistic (Sokal and Rohlf, 1969) comparing $10^{\circ}$ square allele frequencies. By this analysis the samples were not homogeneous at the $5 \%$ level of significance for either 6 -PGD or SDH. In pairwise comparisons, however, no significant difference was found for either 6-PGD or SDH considered alone. Combining data for the two enzymes, Wada found a significant difference in allele frequencies between two of the groups (Antarctic Areas IV and V: on either side of $130^{\circ} \mathrm{E}$ ). He believed this was a more reliable assessment than that presented in Wada and Numachi (1979).

Wada (1984) has also compared minke whale populations off Japan and Korea. A total of 45 whales were sampled from Korean waters and 236 from coastal waters near Japan. Samples were analysed at 15 loci, one of which was polymorphic in both populations. Gene frequency at this locus (ADH-1) was compared by G-statistic for the two populations and found to be significantly different. A chi-square test was used to show that there was no significant deviation from Hardy-Weinberg expectations in either population.

Horwood (1980) applied the genetic distance measures of Rogers (1972), Nei (1972) and Cavalli-Sforza and Edwards (1967) (for full description see Sections 2.4, 2.5)
to the data presented by Wada and Numachi (1979) on minke whales from the Antarctic. By all measures genetic distance between populations from the five Area divisions on the Antarctic whaling grounds were very small (e.g. Nei's $\mathrm{D}=0.001-0.002$ ). Horwood suggested that these distance measures are small compared to subpopulation distances seen in some fish species, and well within the range of the human racial differences described by Nei and Roychoudhury (1974b). As pointed out by Van Beek and Van Biezen (1982), no significant results were found when an a posteriori test with a pair-wise error rate of 0.005 is conducted (overall test significance level of 0.05). The samples from either side of $130^{\circ} \mathrm{E}$ longitude in the Antarctic, however, come close to a significant result. These results should be interpreted with the consideration that very few loci were investigated (three polymorphic loci), which makes distance measures very approximate. Shimura and Numachi applied Nei's (1972) genetic distance measure to twelve species of odontocete cetaceans. Distances between genera were on average 0.213 , distances between families averaged 1.14 , consistent with other mammalian taxonomic divisions (see Section 2.4).

For the population studies on minke whales it was possible to take advantage of the considerable quantity of material made available by the whaling industry. However, as will be described in more detail below, it is more important to examine genetic variation at numerous sites within the genome. The degree of differentiation between minke whale stocks in the Antarctic and North Pacific cannot be fully described by a statistical difference in allele frequencies at a few polymorphic loci. Selective forces could differentiate populations at particular loci, even though the populations continue to interbreed. The potential for describing population differences using theory based on isozyme studies is outlined in the following sections.

### 2.2 Theoretical considerations

Possible explanations for low genetic variability at coding loci for marine and other large mammals are related to the controversy over the evolution and maintenance of variation in natural populations. The neutral gene hypothesis (Kimura, 1968; King and Jukes, 1969) suggests that mutation is the primary force in evolution (see Nei, 1983). By this theory, evolution occurs by the random fixation of neutral or nearly neutral mutations. At any one time, the degree of polymorphism is a consequence of new variation tending towards fixation or elimination by chance. Deleterious alleles are removed by 'purifying' selection. The neutral theory predicts high levels of protein polymorphism in natural populations, and approximate constancy in the rate of amino acid substitutions for each protein (c.f. Zuckerkandl and Pauling, 1965).

Variation can be reduced in natural populations by two stochastic mechanisms, consistent with the neutral hypothesis: the founder effect (a single gravid female colonising a new area; Mayr, 1963; Carson, 1971) and the 'bottleneck' phenomenon (Nei et al., 1975). In either case the population is reduced to a small inbreeding group where loci are readily fixed by genetic drift. From this condition of reduced heterozygosity, the process of gradual accumulation of new mutations would recommence. This process is very slow, approximately the reciprocal of the mutation rate, to reach an equilibrium level (potentially longer than the life of the species) (Nei et al., 1975).

Bonnell and Selander (1974) have suggested that low levels of variability in northern elephant seal populations are a consequence of the decimation of this species by sealers in the last century. The 'bottleneck' size of the population may have been as low as 20 seals. Some cetacean species have also been extensively hunted, especially the slow swimming species (e.g. right and gray whales) heavily hunted in the 18th and 19th centuries, and reduced to very low population levels. However, genetic variation levels are not known for these species.

Allendorf et al. (1979) suggest that the relatively long generation time and low reproductive rate of large-mammal populations would make them especially susceptible to bottleneck effects. The degree to which heterozygosity is reduced during a bottleneck is related to the reduced population size, the duration of the effect limiting population size and the intrinsic growth rate of the species (Nei et al., 1975). However, the period of recovery, which is dependent on the mutation rate, is independent of the intrinsic growth rate. After a period of reduced heterozygosity (determined by the mutation rate), a population should theoretically recover lost variation over a period approximated by the reciprocal of the mutation rate regardless of heterozygosity levels imposed by the bottleneck period (see Fig. 1 in Nei et al., 1975).

An alternative view to the Neutral Theory is that natural selection is the dominant creative force in evolution. Variation in natural populations is thought to be maintained by balancing selection (Karn and Penrose, 1951; Rendel, 1953; see review by Hartl, 1980) and overdominance (e.g. Allison, 1964; Cavalli-Sforza and Bodmer, 1971). Directional selection can limit variation by increasing the representation of one phenotype at the expense of another (see discussion in Falconer, 1981).

The niche-variation model of selection (Levine, 1953; Van Valen, 1965; Levins, 1968; Antonovics, 1971; Hedrick, Ginevan and Ewing, 1976) suggests that specialised organisms occupying a narrowly defined niche will exhibit low genetic variability (see reviews by Bryant, 1974; Nevo et al., 1983). Another theory contrasts 'fine-grained' and 'course-grained' environments (Levins, 1968; Templeton, 1977; Templeton and Rothman, 1978). Fine-grained environments vary seasonally, but are predictable over time and consistent for all individuals. Course-grained environments vary randomly from one generation to the next. By this model, mobile animals with the capacity to select favourable environments and the physiological mechanisms to adapt to variations in the environment (e.g. mammals), will have low levels of heterozygosity in fine-grained environments. Smith and Fujio (1982) surveyed 106 teleost species and found that generalist species adapted to fine-grained environments had lower levels of variation than habitat specialists.

It has been hypothesized that the low levels of heterozygosity found in some fossorial mammals (e.g. see Nevo and Shaw, 1972; Nevo et al., 1974; Nevo, 1978) the American alligator, Alligator mississippiensis (Gartside, Dessauer and Joanen, 1977) and the harp seal, Pagophilus groenlandicus (Lavigne et al., 1978) reflect the relative stability of their environments. Tolliver, Smith and Leftwich (1985) compared genetic diversity in fossorial and terrestrial species within the class Insectivora. They found that fossorial species were less variable, but that terrestrial insectivores were less variable than fossorial rodents (in comparison with data from Selander et al., 1975; Nevo, 1979). The findings of various researchers (e.g. Selander et
al., 1975; Ayala et al., 1975; Patton and Yang, 1977; Nevo, 1978; Tolliver et al., 1985) suggest that exceptions are common (e.g. benthic marine invertebrates exhibit relatively high levels of genetic variation; Ayala et al., 1975) and that a multifactor approach will better predict heterozygosity.

### 2.3 Genetic diversity

A number of formulations have been described to compare genetic diversity between populations within a species. Most of these were developed for the analysis of protein variation, however they are generally adaptable to variation at finer genomic levels, as will be discussed in more detail below (see Nei, 1987). Nei (1975) describes the analysis of gene diversity within sub-divided populations. These measures are related to genotype frequencies only in random mating populations. For a population divided into s sub-populations, the average gene diversity between subpopulations is given by

$$
\mathrm{D}_{\mathrm{ST}}=\mathrm{J}_{\mathrm{S}}-\mathrm{J}_{\mathrm{T}}
$$

where $\mathrm{J}_{\mathrm{S}}$ is the average gene identity (homozygosity) within subpopulations and $J_{T}$ is the average gene identity for the whole population. Gene identity in a given subpopulation is given by

$$
\mathrm{J}_{\mathrm{i}}=\Sigma_{\mathrm{k}} \mathrm{X}^{2}{ }_{\mathrm{ik}}
$$

where $\mathrm{x}_{\mathrm{ik}}$ is the frequency of the k -th allele in the i -th subpopulation. The gene identity for the total population is

$$
\mathrm{J}_{\mathrm{T}}=\Sigma_{\mathrm{k}} \mathrm{x}^{2} \cdot \mathrm{k}
$$

where $\mathrm{x}_{. \mathrm{k}}=\Sigma_{\mathrm{i}} \mathrm{x}_{\mathrm{ik}} / \mathrm{s}$. The gene diversity (heterozygosity) for the total population is $H_{T}=1-J_{T}$, the average gene diversity for subpopulations is $\mathrm{H}_{\mathrm{S}}=1-\mathrm{J}_{\mathrm{S}}$ and $\mathrm{H}_{\mathrm{T}}+$ $H_{S}=D_{\text {ST }}$. Gene diversity between two particular subpopulations (the i -th and j -th populations) is given by

$$
\begin{aligned}
\mathrm{D}_{\mathrm{ij}} & =\left(\mathrm{J}_{\mathrm{i}}+\mathrm{J}_{\mathrm{j}}\right) / 2-\mathrm{J}_{\mathrm{ij}} \\
& =\mathrm{H}_{\mathrm{ij}}-\left(\mathrm{H}_{\mathrm{i}}+\mathrm{H}_{\mathrm{j}}\right) / 2
\end{aligned}
$$

The relative magnitude of gene differentiation among subpopulations (called the coefficient of gene differentiation) is given by

$$
\mathrm{G}_{\mathrm{ST}}=\mathrm{D}_{\mathrm{ST}} / \mathrm{H}_{\mathrm{T}}
$$

However, $G_{S T}$ is dependent on gene diversity. When $H_{T}$ is very small, $\mathrm{G}_{\mathrm{ST}}$ may be artificially large. For this reason Nei also describes a measure that is independent of gene diversity, and estimates the minimum net codon differences between populations. This is called the absolute degree of gene differentiation and is given by

$$
\mathrm{D}_{\mathrm{m}}=\mathrm{sD}_{\mathrm{ST}} /(\mathrm{s}-1)
$$

These formulations lend themselves to hierarchical subdivision so that diversity between colonies within subpopulations or demes, etc. can be described (see examples in Nei, 1975).

A measure that is proportional to the number of alleles in the population, maximum when all alleles are equal in frequency, minimum when there is only a single allele and a converse function of frequencies of alleles was suggested by Lewontin (1972) to characterize genetic diversity: the Shannon information measure:

$$
\mathrm{H}_{\mathrm{o}}=-\Sigma \mathrm{p}_{\mathrm{i}} \mathrm{ln}_{2} \mathrm{p}_{\mathrm{i}}
$$

This measure can also be used to calculate diversity at several levels of gene frequencies. Derivation of $\mathrm{H}_{\mathrm{o}}$ values is facilitated by published tables of $\operatorname{pln}_{2} \mathrm{p}$ (Dolansky and Dolansky, 1952).

Subdivision itself can affect genotype frequencies. If a species is divided into subpopulations where there is random mating, and if gene frequencies differ from population to population, then for the species as a whole, homozygous genotypes will increase at the expense of heterozygotes (Wahlund, 1928). This is known as the Wahlund effect, and it has the same effect on overall heterozygosity as inbreeding.
$\mathrm{G}_{\mathrm{ST}}$ can be regarded as an extension of Wright's correlation between two gametes drawn at random from each subpopulation, the F-statistic $\mathrm{F}_{\text {ST }}$ (Wright quoted in Nei, 1975). This is an estimation of the fixation index from a group of sample populations and is given by the actual gene frequency variance divided by the limiting variance for each allele:

$$
\mathrm{F}_{\mathrm{ST}}=\mathrm{s}_{\mathrm{q}(\mathrm{ST})}{ }^{2} / \mathrm{q}_{\mathrm{T}}\left(1-\mathrm{q}_{\mathrm{T}}\right)
$$

However, Wright's (1965) (and Cockerham's, 1969; 1973) application of the F-statistic was devised in terms of neutral genes. Further there is an assumption that the number of subpopulations is infinitely large. The principle distinction between Wright's and Nei's formulations is that Wright defines the F-statistic as a correlation between uniting gametes, while Nei compares observed and expected heterozygosities. Nei (1977) re-defines the application of F-statistics to this problem as a function of heterozygosities. So defined it is independent of the number of subpopulations or alleles involved and can be applied whether or not there is selection. As defined in Nei (1977), $\mathrm{F}_{\mathrm{ST}}$ is identical to $\mathrm{G}_{\mathrm{ST}}$. As with the measures described in Nei (1975), $\mathrm{G}_{\text {ST }}$ and $\mathrm{F}_{\text {ST }}$ are not related to the frequency of heterozygotes except in populations that are consistent with the Hardy-Weinberg rule.
Nei (1975) has applied $G_{S T}$ and $D_{m}$ to a number of studies on protein polymorphisms in regional populations and races. Nei and Roychoudhury (1982) looked at 62 protein loci in the three principle human races (Caucasoid, Negroid and Mongoloid). The minimum net codon differences between the three races were estimated to be 0.0195 per locus ( $\mathrm{D}_{\mathrm{m}}$ ). $\mathrm{G}_{\mathrm{ST}}$ was estimated to be 0.088 , which means that $8.8 \%$ of the total gene diversity can be attributed to genetic differences between the races. Clearly variation within races accounts for most of the genetic diversity.
Nei's (1975) analysis of data from a study of 37 villages of the Yanomama Indians (Weitkamp et al., 1972) gave a comparable value of $\mathrm{G}_{\mathrm{ST}}=0.069$. Applying this technique to other species gave values ranging from 0.072 for 4 populations of horseshoe crab (Selander et al., 1970) to 0.284 for 4 populations of Lycopodium lucidulum (Levin and Crepet, 1973). A study on 9 populations of Dipodomys ordii (Johnson and Selander, 1971) provides an example of how a low value for $H_{S}$ can influence $G_{S T}$. $G_{S T}$ for this study works out at 0.674 , but $\mathrm{H}_{\mathrm{S}}$ is quite low ( 0.012 ), and $\mathrm{D}_{\mathrm{m}}$ is about the same as for the study on Lycopodium lucidulum.

### 2.4 Genetic distance

Genetic distance is a measure of gene diversity between populations expressed as a function of genotype frequency (Nei, 1972). Various authors have suggested measures of
genetic distance (Sanghvi, 1953; Prevosti, 1955; Sokal and Sneath, 1963; Cavalli-Sforza and Edwards, 1967; Balakrishnan and Sanghvi, 1968; Hendrick, 1971; Nei, 1972; Rogers, 1972). Nei (1972), Rogers (1972) and Wright (1978) discuss the relative benefits of the various methods.

The simplest measure of distance for a single locus is a measure proposed by Prevosti (1955). For multiple loci it is the arithmetic mean of half the sum of the absolute differences between allelic frequencies:

$$
\mathrm{D}=0.5 \Sigma\left|\mathrm{q}_{\mathrm{x}}-\mathrm{q}_{\mathrm{y}}\right|
$$

Wright (1978) describes the theoretical problem with this model: that equal weight is given to frequency differences throughout the range from 0 to 1 , and suggests a transformation of scale. However, after transformation two populations with no alleles in common will no longer give a distance value of 1.0 , if one of the populations carries three or more alleles at a locus.

Rogers (1972) proposed a distance measure based on an extension of the Pythagorean theorem. For a single locus, the genetic distance is given by

$$
\mathrm{D}_{(\mathrm{XY})}=\left(0.5 \Sigma\left(\mathrm{q}_{\mathrm{xi}}-\mathrm{q}_{\mathrm{yi}}\right)^{2}\right)^{1 / 2}
$$

The distance with respect to multiple loci is the arithmetic mean of the coefficients for the separate loci. This measure gives a value of zero for populations with identical alleles, but the value is less than one for completely dissimilar populations, if there are multiple alleles. The distance measure proposed by Cavalli-Sforza and Edwards (1967) is equivalent except that they take the square roots of the allelic frequencies. This solves the problem encountered by Rogers for multi-allelic cases.

Rogers (1972) applied the formulations developed by Sokal and Sneath (1963), Cavalli-Sforza and Edwards (1967), Hendrick (1971), and his own measure to data from a Danish house mouse population studied by Selander, Hunt and Yang (1969). From this he found that the different approaches produce very similar results, although he suggests that there are difficulties with some of the methods. For example, by Sokal and Sneath's measure of similarity, two populations can have identical allele frequencies, but not have the maximum similarity value; slightly dissimilar populations may have a higher value (Rogers, 1972). In general, a method well suited to comparing genetic distance between populations within a species is that proposed by Nei (1971; 1972; 1975; 1978).

The accuracy of Nei's formulations is limited primarily by the proportion of variation that can be detected by electrophoresis, and any variation in the rate of nucleotide substitution at different loci. Other formulations are limited by these same problems. Nei suggests that his measure of genetic distance has the advantages of measuring the accumulated number of gene substitutions per locus, and having a linear relation to evolutionary time (assuming a constant rate of nuclear substitution). Nei (1975) defines three measures of genetic distance, the minimum, standard and maximum measures.

As described above, the probability of identity of two genes chosen at random in a population is

$$
\mathrm{j}_{\mathrm{x}}=\Sigma \mathrm{x}^{2}{ }_{\mathrm{i}}
$$

where $x_{i}$ is the frequency of the $i$-th allele. The probability of identity of two genes chosen at random, one from each of two populations X and Y is

$$
j_{x y}=\sum x_{i} y_{i}
$$

The gene frequencies used are those observed in the population. No assumptions about selection, mutation or migration are required. A large number of loci should be investigated. The arithmetic means over all loci (including monomorphic loci) for $\mathrm{j}_{\mathrm{x}}, \mathrm{j}_{\mathrm{y}}$ and $\mathrm{j}_{\mathrm{xy}}$ are designated $\mathrm{J}_{\mathrm{X}}, \mathrm{J}_{\mathrm{Y}}$ and $\mathrm{J}_{\mathbf{X Y}}$. The normalised identity of genes between populations X and Y is given by

$$
\mathrm{I}=\mathrm{J}_{\mathbf{X Y}} / \sqrt{\mathrm{J}_{\mathbf{X}} \mathrm{J}_{\mathbf{Y}}}
$$

If individual codon changes are independent, the mean number of net codon differences, the standard genetic distance (Nei, 1972), is given by

$$
\mathrm{D}=-\log _{\mathrm{e}} \mathrm{I}
$$

Minimum estimates of codon differences can be determined by

$$
\begin{aligned}
\mathrm{D}_{\mathrm{X}(\mathrm{~m})} & =1-\mathrm{J}_{\mathrm{X}} \\
\mathrm{D}_{\mathrm{Y}(\mathrm{~m})} & =1-\mathrm{J}_{\mathrm{Y}} \\
\mathrm{D}_{\mathrm{XY}(\mathrm{~m})} & =1-\mathrm{J}_{\mathrm{XY}}
\end{aligned}
$$

The minimum genetic distance (Nei, 1975) is given by

$$
\mathrm{D}_{\mathrm{m}}=\mathrm{D}_{\mathrm{XY}(\mathrm{~m})}-\left(\mathrm{D}_{\mathrm{X}(\mathrm{~m})}+\mathrm{D}_{\mathrm{Y}(\mathrm{~m})}\right) / 2
$$

When there are only two populations this is the same as the interpopulational gene diversity described above ( $\mathrm{D}_{\mathrm{m}}$ ). The main drawback with this measure is that the estimates of codon differences are not additive so that $\mathrm{D}_{\mathrm{m}}$ will greatly underestimate the number of net codon differences when $\mathrm{D}_{\mathrm{XY}(\mathrm{m})}$ is large. If the rate of codon changes vary from locus to locus (as is the case; see discussion in Section 1.4), D will also be an underestimate. In this case genetic distance can be estimated with the same formulation as for D , except that $\mathrm{J}_{\mathrm{XY}}, \mathrm{J}_{\mathrm{X}}$ and $\mathrm{J}_{\mathrm{Y}}$ are computed as the geometric means of $j_{x y}$, $j_{x}$ and $j_{y}$. This is designated the maximum genetic distance, D' (Nei, 1975). Sampling errors of gene frequencies can greatly inflate this measure however, and if there is even one locus where there is no common allele between two populations, D' will be infinitely large. Nei (1975) notes that for most practical applications of these measures the difference between $\mathrm{D}_{\mathrm{m}}$, D and $\mathrm{D}^{\prime}$ is very small. Also $\mathrm{D}_{\mathrm{m}}<\mathrm{D}<\mathrm{D}$ ' except when these quantities are very small. Standard errors for these formulations have been computed by Nei and Roychoudhury (1974a, see below).

Table 2
Variation in genetic distance D, after Nei, 1987

| Taxa/species | Taxa | Loci | D | References |
| :---: | :---: | :---: | :---: | :---: |
| Local races |  |  |  |  |
| Fish ${ }^{1}$ | 4 | 33 | 0.000-0.003 | Bush \& Crabtree, 1982 |
| Red deer | 4 | 34 | 0.016 | Gyllensten et al., 1983 |
| Pocket gophers | 10 | 31 | 0.004-0.262 | Nevo et al., 1974 |
| Species |  |  |  |  |
| Galapagos finches | 6 | 27 | 0. 004-0.065 | Avise et al., 1980 |
| Salamanders ${ }^{2}$ | 26 | 29 | 0.18-3.00 | Highton \& Larsen, 1979 |
| Lizards ${ }^{3}$ | 4 | 23 | 1.32-1.75 | Webster et al., 1972 |
| Genera |  |  |  |  |
| Galapagos finches | 5 | 27 | $0.04-0.14$ | Yang \& Patton, 1981 |
| Fish ${ }^{4}$ | 3 | 31 | 0.47-1.30 | Ward \& Galleguillos, <br> 1978 |
| Families |  |  |  |  |
| Man-chimpanzee | 2 | 42 | 0.62 | King \& Wilson, 1975 |
| 1 Catostomos sp. |  |  | nalis sp. |  |
| 2 Pibthodon sp. |  |  | leuronectida |  |

Nei (1975) has applied his measure of standard genetic distance ( $\mathrm{D}=-\ln \mathrm{I}$ ) to a number of studies on populations within species, sub-species and higher taxonomic divisions. Genetic distance between races (or populations) was always less than a few percent. Nine populations of the kangaroo rat (Dipodomys ordii) (Johnson and Selander, 1971) showed the greatest variation between pairs of populations with D ranging from 0.000 to 0.058 . Genetic distance between human races varied from 0.011 to 0.019 (Nei and Roychoudhury, 1974b). This is equivalent to a duration of 55,000 to 95,000 years reproductive isolation (using the formulation $t=5 \times 10^{6} \mathrm{D}$, see below). Seven studies comparing subspecies showed a range of $D$ values from 0.004 to 0.351 . The majority were about an order of magnitude higher than genetic distances between races. At all levels there is considerable variation in the estimates of D (Table 2). However on average higher taxanomic divisions had higher values.

### 2.5 Variance of heterozygosity and genetic distance

If the heterozygosity at a given locus in a population is given by $h$, and the number of loci examined is $r$, then the average heterozygosity over all loci is given by

$$
\mathrm{H}=\Sigma \mathrm{h}_{\mathbf{k}} / \mathrm{r}
$$

where $h_{k}$ is the heterozygosity at the $k$-th locus. The expected variance of $h$ is estimated by

$$
V(\mathrm{~h})=\Sigma\left(\mathrm{h}_{\mathrm{k}}-\mathrm{H}\right)^{2} /(\mathrm{r}-1)
$$

The sampling variance of H is given by

$$
V(H)=V(h) / r
$$

This assumes that h's at different loci are independent, which is generally the case unless there is linkage disequilibria. Nei and Roychoudhury (1974a) describe two measures of sampling variance: the inter-locus and the intra-locus variances. The inter-locus variance is determined by diverse evolutionary forces, and is usually very difficult to quantify. The intra-locus variance is dependent on the sample size and the gene frequencies of the locus studied. This measure is used to compute the standard errors of heterozygosity and genetic distance, and to estimate the magnitude of inter-locus variance. The total variance is equal to the sum of the component variances so that

$$
V(h)=V_{g}(h)+V_{s}(h)
$$

where $V_{g}(h)$ is the inter-locus variance and $V_{s}(h)$ is the intra-locus variance. The intra-locus variance can be estimated by

$$
\mathrm{V}_{\mathrm{s}}(\mathrm{~h})=\Sigma \mathrm{V}_{\mathrm{s}}\left(\mathrm{~h}_{\mathrm{k}}\right) / \mathrm{r}
$$

where $r$ is the number of loci studied. The inter-locus variance is estimated by

$$
\mathrm{V}_{\mathrm{g}}=(\mathrm{n}-1)^{2} \mathrm{~V}_{\mathrm{h}}(\mathrm{~h}) / \mathrm{n}^{2}
$$

where n is the number of genes and $\mathrm{V}_{\mathrm{h}}$ is the variance of homozygosity and heterozygosity among loci. A similar formulation can be derived for the variance of the minimum genetic distance, thus

$$
V(d)=\Sigma\left(d-D_{m}\right)^{2 /(r-1)}
$$

where $d$ is the minimum distance measure per locus. The formulations for $D$ and $D^{\prime}$ are more involved. The mathematical argument for an asymptotic approximation is presented in Nei and Roychoudhury (1974a). Further
discussion on sampling error in the determination of genetic heterogeneity within and between populations can be found in Workman and Niswander (1970).

These variance measures are affected by deviation from the Hardy-Weinberg equilibrium, and dominance. Dominance will tend to increase the variance, although the effect is small unless the frequency of recessive genes is very small. Inbreeding is expected to increase the variance.

When planning a study on heterogeneity and genetic distance between sample populations, the number of loci and individuals that should be investigated can be approximated by minimizing the sampling variance (Nei and Roychoudhury, 1974a).

$$
\mathrm{V}(\mathrm{H})=\left(\mathrm{V}_{\mathrm{g}}(\mathrm{~h})+\mathrm{V}_{\mathrm{s}}(\mathrm{~h})\right) / \mathrm{r}
$$

If the total number of genes to be studied is held constant (rn), then clearly $\mathrm{V}(\mathrm{H})$ can be minimised by maximizing r (the number of loci studied). Nei and Roychoudhury illustrate this point by analysing data obtained by Avise and Selander (1972). This was a study on three cave and nine surface populations of a characid fish species. Sample variances were computed for comparisons of populations of different sizes and for different numbers of loci. For the purpose of estimating average heterozygosity or genetic distance, they suggest that as few as 20 individuals per locus would be sufficient as long as a large number of loci (say 30 to 70) are investigated. However if the number of individuals is too small, the bias of the heterozygosity estimate becomes large. Nei (1978) presents a detailed analysis of the affect of small sample sizes on his measures of heterozygosity and genetic distance. The general conclusions were that few individuals need be studied when genetic distance is fairly high, and when heterozygosity levels are low (providing numerous loci were analysed). In the converse cases, and when the number of loci that can be investigated is limiting, a larger number of individuals will improve the results.

### 2.6 Genetic distance and evolutionary time

The estimates of genetic distance given above are based on codon differences per locus. This means that a large number of loci must be examined to achieve an approximation that is close to the real value. When comparing local populations of the same species, deviations from the true value are expected to be upward when only a few loci are available for study. One reason for this is that monomorphic loci in these populations will usually have the same allele (Nei, 1975). In any case, these measures are useful as estimates of relative distance because they do not depend on assumptions about evolutionary forces. Estimates of genetic distance can also be achieved based on the rate of gene substitution per locus per year (a).

The normalized gene identity between two isolated populations under mutation pressure and independent of selection can be given by

$$
\mathrm{I}=\mathrm{I}_{\mathrm{o}} \mathrm{e}^{-2 \mathrm{at}}
$$

where $t$ is the time since reproductive isolation (Nei and Feldman, 1972). $I_{o}$ is the initial gene identity, and should be close to one. No appreciable gene differentiation would be expected as long as there is migration between populations (to be discussed in more detail below). According to this formulation, the gene identity should decrease exponentially as $t$ increases. Two assumptions are necessary for this definition. First, the two populations
under comparison must be in equilibrium with respect to random genetic drift, mutation and selection. This means that the average gene identities ( $\mathrm{J}_{\mathrm{X}}$ and $\mathrm{J}_{\mathrm{Y}}$ ) will remain constant. Second, the rate of gene substitution per locus per year (a) should remain constant. Substituting this version of I into the formulation for D gives approximately 2at.

The main problem with this estimate is the crudeness of the approximation of a. By electrophoretic analysis only a proportion of the existing variation can be detected. Let c be the proportion of amino acid differences detectable by electrophoresis, then

$$
\begin{aligned}
\mathrm{I} & =\mathrm{e}^{-2 \mathrm{cat}} \\
\text { so }, \mathrm{D} & =2 \mathrm{cat}
\end{aligned}
$$

The number of codon differences (2at) can then be estimated by D/c. Nei and Chakraborty (1973) suggest that this estimation is only applicable when 2 at $<2$. This is because the chance of accumulating charge differences in amino acid combinations that cancel out (and are therefore undetectable) is expected to increase with time since divergence.

The above formulation for D can be rewritten to give an estimate of the time since two populations became reproductively isolated, assuming a constant rate of gene substitution (Nei, 1971): $\mathfrak{t}=2 \mathrm{aD}$. If a is estimated to be roughly $10^{-7}$ for proteins detectable by electrophoresis, then $\mathrm{t}=5 \times 10^{6} \mathrm{D}$. Nei (1971) suggests that this estimate is appropriate for values of $D<1$. For large values of $D$ and when a varies among loci, $t$ is an underestimate. As noted above, gene substitution rates do vary considerably in different regions of the genome, and for different protein loci. For example, King (1973) suggested an order of magnitude difference in the substitution rates for intracellular versus extracellular proteins. Therefore measures of t will be underestimates by this method.

### 2.7 Effect of migration on population diversity

Surprisingly little mixing is necessary to overcome the effects of genetic drift and maintain genetic homogeneity between populations. Crow and Kimura (1970) describe the conditions for the establishment of an equilibrium between migration and random drift. In a random mating population the probability that two gametes will have identical genes is $1 / 2 \mathrm{~N}$. The chance that two gametes have different parental genes is $1-1 / 2 \mathrm{~N}$ ( 2 N is the number of genes at a given locus in a population of N diploid parents). The inbreeding coefficient is given by

$$
\mathrm{f}_{\mathrm{t}}=1 / 2 \mathrm{~N}+(1-1 / 2 \mathrm{~N}) \mathrm{f}_{\mathrm{t}-1}
$$

where $f_{t-1}$ is the inbreeding coefficient for an average individual in the previous generation. If we consider a group of subpopulations with a migration rate between them of $M$, the probability that neither of the two uniting genes will be displaced by migrant gene is $(1-\mathrm{M})^{2}$. The increase in autozygosity of a subpopulation is given by

$$
\mathrm{f}_{\mathrm{t}}=\left(1 / 2 \mathrm{~N}_{\mathrm{e}}+\left(1-1 / 2 \mathrm{~N}_{\mathrm{e}}\right) \mathrm{f}_{\mathrm{t}-1}\right)(1-\mathrm{M})^{2}
$$

where $\mathrm{N}_{\mathrm{e}}$ is the effective (reproducing) number in the subpopulation. At equilibrium $f_{t}=f_{t-1}=f$, and when $M$ is small $\mathrm{M}^{2}$ can be neglected, so that

$$
\mathrm{f}=1 / 4 \mathbf{N}_{\mathrm{e}} \mathbf{M}+1
$$

This means that if $M$ is very much less than $1 / 4 N_{e}$, then $f$ will be large and the populations will tend to diverge. However, if $M$ is larger than $1 / 4 N_{e}$, then the
subpopulations are effectively a single panmictic unit. In practical terms, if one or more reproductively active individual migrates between sub-populations per generation, then there will be little local differentiation. The effect will be less pronounced in species that tend to disperse over a short range, because neighboring subpopulations will tend to be genetically similar. Further differentiation could result from selection pressure (see below) or a variety of DNA turnover mechanisms (see Section 1.3). In this case a higher level of exchange would be necessary to eliminate genetic differentiation between populations.

Considering only the effects of genetic drift, Nei and Feldman (1972) develop a formulation for the effect of migration on the normalized identity of genes between two populations. If $M=m_{1}+m_{2}$, where $m_{1}$ and $m_{2}$ stand for the migration rates from populations 1 and 2 respectively, then

$$
\mathrm{I}=\mathrm{M} /(\mathrm{M}+2 \mathrm{u})
$$

where u is the mutation rate per locus per generation. The quantity 2 u is very small, so ' $I$ ' will be very nearly 1 unless M is very small. This means that genetic distance between populations cannot be large unless migration rates are very low.

Given a computed measure of genetic distance between two populations, it is possible to estimate the maximum possible migration rate that could have occurred (Nei, 1975). Assuming that the genetic distance between the populations has reached a steady-state value, then $\mathrm{I}=$ $\exp (-\mathrm{D})=\mathrm{m} /(\mathrm{m}+\mathrm{u})$, where m is the maximum migration rate and $u$ is the mutation rate per locus per generation. Therefore,

$$
\mathrm{m}=\mathrm{u} \exp (-\mathrm{D}) / 1-\exp (-\mathrm{D})
$$

Comparing human races, assuming a mutation rate of $2 \times$ $10^{-6}$ per generation, Nei (1975) derives a maximum migration rate of $1 \times 10^{-4}$ per generation between Caucasoids and Negroids, and $2 \times 10^{-4}$ between Caucasoids and Mongoloids.

Wright (1940; 1951) discusses the interaction between migration and selection for his island model. Sub-populations are thought of as being isolated, but with limited exchange of individuals. The mathematical argument is summarised by Crow and Kimura (1970), and will not be detailed here. By this model the populations will differentiate only when the selective force is much larger than the influence due to migration. When selective and migration forces are equal, a gene that is favoured by local selection will have a frequency equal to the square root of the average allele frequency for the entire population.

### 2.8 Gene frequency comparisons by chi-square and F-statistics

The models discussed above have been primarily concerned with multi-loci comparisons. Often it is desirable to investigate variation between populations at a specific locus. This is usually done by chi-squared analysis, or by the F-statistic. Workman and Niswander (1970) discuss the relationship between the two statistics. The chi-squared statistic is used to describe heterogeneity of gene frequencies between populations by pairwise comparisons in a contingency table analysis. This relationship is given by

$$
\chi^{2}=(2 N) s_{p}^{2} \sqrt{p q}
$$

where $N$ is the sample size and $s_{p}{ }^{2}$ is the weighted variance at the p allele.

$$
\begin{gathered}
\mathrm{s}_{\mathrm{p}}^{2}=\Sigma\left(\mathrm{N}_{\mathrm{i}} / \mathrm{N}\right) \mathrm{p}^{2}{ }_{\mathrm{i}}-\overline{\mathrm{p}}^{2} \\
\mathrm{~F}_{\mathrm{ST}}=\mathrm{s}_{\mathrm{p}}^{2} / \overline{\mathrm{pq}}, \text { so that } \\
\mathrm{F}_{\mathrm{ST}}=\mathrm{X}^{2} / 2 \mathrm{~N}
\end{gathered}
$$

where $p=\Sigma\left(N_{i} / N\right) p_{i}$. For a $k$-allelic locus,

$$
\mathrm{X}^{2}=2 \mathrm{~N}(\mathrm{k}-1) \mathrm{F}_{\mathrm{ST}}
$$

Therefore, $\mathrm{F}_{\mathrm{ST}}$ is related to the chi-square statistic as a simple function of the sample size. An estimate of $F_{S T}$ is directly dependent on the significance of the chi-square statistic. Both chi-square and F -statistics are used to describe observed deviations from theoretical expectations.

For comparing observed genotypic proportions with those expected in a Hardy-Weinberg population, chi-square and $\mathbf{F}$ are related in the following way

$$
\mathrm{X}^{2}=\mathrm{F}^{2} \mathrm{~N}
$$

where F can be estimated by $\mathrm{F}=1-\mathrm{H} / 2 \mathrm{pq}$ and N is the sample size. H is the observed level of heterozygosity. Presented in terms of $F$ this gives

$$
F=\left(X^{2} / N\right)^{1 / 2}
$$

which is the coefficient of contingency for a 2 X 2 contingency table (see Kendall and Stuart, 1961). Various authors (e.g. Workman, 1969; Neel et al., 1964) suggest that deviations from the Hardy-Weinberg law will rarely be significant by this test, even in small inbred populations. When a deviation is significant, this does not imply anything about selective forces or distinguish random from assortive mating. It can only suggest that there were no sampling biases or biological influences sufficient to cause a significant deviation from theoretical expectations (Workman and Niswander, 1970).

Unless genetic differences are quite extreme, a large sample size is required to distinguish populations at a single locus. This relationship has been described by Sharp (1976) for comparisons by the chi-square test. A minimum limiting factor is that expected frequencies should have at least 5 events per class. An a priori decision should be made on the following factors: (1) the minimum difference in allele frequencies that will be counted as significantly different (d); (2) the significance level for a type one error (rejection of the null hypothesis); (3) the significance level for a type two error (the probability that a false null hypothesis will be accepted); (4) the expected range of allele frequency values. If $P$ is the relative frequency at a particular allele, and $n_{1}$ and $n_{2}$ are the sample sizes from populations 1 and $2\left(\mathrm{n}_{1}+\mathrm{n}_{2}=\mathrm{N}\right)$, then

$$
l=\mathrm{d}^{2} \mathrm{n}_{1} \mathrm{n}_{2} / \mathrm{P}(1-\mathrm{P}) \mathrm{N}
$$

where $l$ is the non-centrality parameter for the non-central chi-square distribution (see Abramowitz and Stegun, 1964). If sample sizes are equal then this is equal to

$$
\mathrm{d}^{2} \mathrm{~N} / 4(\mathrm{P})(1-\mathrm{P})
$$

In biochemical studies the number of individuals to be studied is $\mathrm{N} / 2$. Therefore the number of individuals to be sampled should be

$$
\mathrm{N}_{\mathrm{i}}=2(\mathrm{P})(1-\mathrm{P})(\mathrm{l}) / \mathrm{d}^{2}
$$

Given a fixed type one error (usually 0.05 ) the number of individuals that would allow differentiation of two populations can be determined for different values of $B$
(type two errors) and $d$ (permitted magnitude of the difference between gene frequencies). Numerical examples are presented in Sharp (1976).

A computer program, BIOSYS-1*, has been developed that tests data sets for most measures of genetic variability (Swofford and Selander, 1981). This includes measures of heterozygosity, conformance to Hardy-Weinberg expectations, heterogeneity by Wright's F-statistic and chi-square analysis, and the similarity and distance measures of Rogers, Nei, Prevosti, and Cavalli-Sforza and Edwards. The data can also be analysed by cluster analysis to produce dendrograms and by the method of Farris (1970) to produce Wagner trees.

### 2.9 Kinship assessment

In many cases, especially with the social odontocetes, it would be possible to interpret important details about population mixing and dynamics if breeding system and kinship within social groups were known. Genetic markers revealed by allozyme analysis can be used for these determinations (although far more powerful techniques are available; see discussion below).

Tests for paternity are possible by exclusion analysis. This technique has been applied to a variety of taxa (e.g. bats - McCracken and Bradbury, 1977; 1981; Porter and McCracken, 1983; rodents - Foltz, 1981; Hanken and Sherman, 1981; lagomorphs - Daly, 1981; primates Smith, 1980; insects - Pamilo, 1982; McCauley and O'Donnell, 1984). If a female and her offspring can be identified and their allelic patterns determined, then potential fathers can be screened. Males not possessing the necessary alleles can be eliminated as possible fathers. The identification of cow/calf pairs and potential fathers will be most practical in cetacean species that have been the subjects of long-term photo-identification studies. Individual identification has been possible by this method for a number of species (e.g. right whales: Payne, 1972; killer whales: Bigg, MacAskie and Ellis, 1976; humpback whales: Katona et al., 1979; minke whales: Dorsey, Hoelzel and Stern, 1982). Females and their calves keep in close proximity for varying periods after birth in all species.

Various statistical methods have been employed to calculate the probability of 'non-paternity' (McCracken and Bradbury, 1977; 1981) and the 'likelihood of paternity' (Foltz, 1981; Foltz and Hoogland, 1981) from allozyme data. It is also possible to identify a skew in male reproductive success by examining whether the distribution of paternal allele frequencies in offspring differs from allele frequencies in the adult male population (see McCracken and Bradbury, 1977; 1981).

Kinship among and within social groups has been measured primarily by two methods: by comparing genetic heterogeneity between groups; and by estimating relatedness among individuals within groups. By the first method, an assumption is made that kin groups should show a non-random distribution of allele frequencies, with greater homogeneity within than between kin-groups (see McCracken and Bradbury, 1977; 1981; Patton and Feder, 1981; Daly, 1981). Social groups can be compared at a single locus by the G-test for heterogeneity (Sokal and Rolf, 1969) for evidence of a non-random distribution among groups. Use of the F-statistic (see application by

[^7]Schwartz and Armitage, 1980) allows further partitioning of genetic variance into within group, among group, and among population components. A multi-locus approach has been applied by Wilkinson (1985). This involves the use of discriminant function analysis as described by Smouse, Speilman and Park (1982).

The estimation of relatedness within groups is usually achieved by regression analysis on gene frequencies. Theoretically this can give an average coefficient of relatedness for groups of two or more individuals (Pamilo and Crozier, 1982; Pamilo, 1984; see applications by Pamilo, 1982; Ward, 1983). The standard error of these estimates is determined by the relative frequency of alleles, the number of individuals in a group and the number of groups used in the regression (Pamilo and Crozier, 1982; Pamilo, 1984). Wilkinson and McCracken (In Press) have calculated by simulation study that this technique is most precise when an average regression coefficient is determined from analysis of a number of independent loci.

These estimates of genetic relatedness are based on genetic similarity that could result from either assortive dispersal or common descent. To distinguish these two possibilities, it is necessary to obtain information from natural populations on dispersal patterns.

### 2.10 Utility of protein variation analysis

The main limitation to comparing populations by isozyme studies is the low level of variation relative to other regions of the genome. The advantages are related to the fact that speculation about function is less tentative than for some of the non-transcribed regions. This is important to investigations on the evolutionary role of natural selection. However, to describe genetic distance, degree of reproductive isolation and genealogical relationships within populations, it is more important to maximise the amount of detectable variation. This is more effectively accomplished with analytical techniques that examine DNA variation directly, especially within hypervariable regions of the nuclear genome.

Enzyme polymorphisms at a given set of loci can provide only a statistical distinction between populations. That is, although the means for the two populations may differ, there is generally considerable overlap between the two distributions of multilocus genotypes. Given some a priori criteria for separating populations and sufficient variation, it is possible to use genetic distance measures to distinguish breeding stocks. However, when sorting stocks from a mixed assemblage, it is not possible to use allele frequencies to classify individuals chosen at random, although a maximum likelihood method can be used to estimate the composition of the mixture (Pella and Milner, 1987). One solution is to look for a unique allele, or genetic 'marker', that is indicative of a given population. Unfortunately enzyme variation in marine mammals is fairly low and such markers are rare. However, the characteristics of variation in gene families such as the rDNA region and in mtDNA, as described below, are well suited to this kind of comparison (see Sections 3 and 4.2). Sorting mixed stocks by genetic markers will be facilitated by using multiple marker systems (e.g. both rDNA and mtDNA) and sampling associating animals known to be from the same stock (e.g. cow/calf pairs).

Nei (1987) describes the modification of formulations on variation and genetic distance to accommodate the greater resolution possible through examining DNA sequence more directly. He defines the DNA sequence equivalent of
a gene as a 'nucleon' and variations in the sequence (alleles) as haplotypes or nucleomorphs. Haplotype polymorphisms can be computed in the same way as allele frequencies. The following sections describe the potential for these studies in more detail.

## 3. DNA VARIATION: THE MITOCHONDRIAL GENOME

### 3.1 Variation in mitochondrial DNA

The mitochondrial genome is a circular, double-stranded molecule ranging in size from 15.7 kilobases to 19.5 kilobases in multicellular animals (e.g. Fauron and Wolstenholme, 1976; Brown, 1983). It is functionally different from the nuclear genome in a number of respects. For example, replication is asymmetric, unidirectional and continuous, requiring far fewer enzymes than the symmetric, bidirectional, discontinuous replication of nuclear DNA. The gene content is apparently invariant across all metazoans studied so far (primarily vertebrates and Drosophila spp.) and limited to 13 proteins, 2 ribosomal RNAs and 22 transfer RNAs (see review by Brown, 1985). Replication and translation are initiated in the 'control region' (Brown, Shine and Goodman, 1978; Gillum and Clayton, 1978) where variable non-transcribed regions are also concentrated.

Mitochondrial sequence variation has been investigated by a number of methods including buoyant density shift and thermal stability analysis of homo- and heteroduplexes (see Borst, 1972). A great improvement in resolution was achieved when restriction enzyme analysis was applied (Brown and Vinograd, 1974; Robberson, Clayton and Morrow, 1974). Restriction enzymes recognize particular DNA sequences, usually $4-6$ bases long, and cleave double stranded DNA at that point. The use of ten enzymes, each recognizing a nucleotide sequence of four bases, can resolve mitochondrial genomes that differ by less than $0.05 \%$ (Wilson et al., 1985). Vertebrate mtDNA is cut into about 60 segments by a typical four-base enzyme (e.g. Brown, 1980; Ferris, Sage and Wilson, 1982). Since the genome is circular, an enzyme that cuts the mtDNA $n$ times will produce $n$ fragments. If it is assumed that all nucleotides are randomly distributed in the genome, then the expected frequency of a given restriction site can be estimated by

$$
\mathrm{a}=(\mathrm{g} / 2)^{\mathrm{r} 1}((1-\mathrm{g}) / 2)^{\mathrm{r} 2}
$$

where $r 1$ is the number of guanines $(G)$ and cytosines $(C)$, and r 2 the number of adenines $(\mathrm{A})$ and thymines $(\mathrm{T})$ in the restriction site, while $g$ is the percentage $G+C$ content in the genome ( Nei and $\mathrm{Li}, 1979$ ). For example, if $\mathrm{g}=0.5$ and the size of the mitochondrial genome is estimated at 16,000 base pairs, then the expected frequency of restriction sites for the enzyme EcoRI (which cuts at GAATTC) will be $0.03 \%$ (producing 3.9 restriction fragments). The enzyme AluI, which cuts at the four base sequence: AGCT, should produce 62.5 fragments (see Lansman et al., 1981). DNA sequencing (see Maxam and Gilbert, 1977; Sanger, Nicklen and Coulson, 1977) provides even greater resolution, however it is at present too expensive and time consuming to be a practical alternative for population studies.

Sequence changes in animal mitochondrial genomes are of four principle types: sequence rearrangements, additions, deletions and nucleotide substitutions (see Brown, 1985). Overall substitution rates for the
mitochondrial genome have been estimated to be 5 to 10 times greater than in 'single-copy' nuclear DNA (Brown, George and Wilson, 1979). The lowest mtDNA substitution rates are in the tRNA and rRNA genes (Brown et al., 1982; Cann, Brown and Wilson, 1984). The mtDNA protein genes evolve at about twice that rate (which can be up to two orders of magnitude higher than their nuclear counterparts - Brown et al., 1982; Cann et al., 1984). Rates vary considerably between proteins and at a given protein among different species (e.g. Brown and Simpson, 1982).

There are noncoding sequences in the mitochondrial genome, although proportionally far fewer than in the nuclear genome. Most of these sequences are found immediately adjacent to structural genes and are quite small (usually 5 base pairs or less in vertebrates - e.g. Andersen et al., 1981; Bibb et al., 1981). Substitution at these sites occurs at about the same frequency as for synonymous third position codon sites in protein genes (which evolve at 3-4 times the rate of non-synonymous codon sites - e.g. Cann and Wilson, 1983).

The most variable part of the mitochondrial genome is the region where replication begins (e.g. Walberg and Clayton, 1981; Chang and Clayton, 1985). It is estimated that the size of this region varies among animal species from about 200 to 4,100 base pairs (Brown, 1985). The substitution rate in the control region is estimated to be 2.8 (Cann et al., 1984) to 5 (Aquadro and Greenberg, 1982) times the rate found in the remainder of the genome. There are three conserved blocks near the promoter sequences, one of which has been associated with a function (Chang and Clayton, 1985). However, these represent a very small portion of the total region.

In general, the mitochondrial genome is considerably more variable than nuclear DNA. As discussed above, when measuring heterogeneity within or genetic distance between populations, accuracy is greatly enhanced by increasing the number of loci investigated. Sequencing studies have confirmed earlier suggestions that most polymorphism in mammalian mtDNA results from base substitution (e.g. Greenberg, Newbold and Sugino, 1983). This would allow restriction polymorphisms to be interpreted as changes in the nucleotide sequence. Enzyme variation can only indicate a change somewhere in a long sequence that codes for a protein. Therefore, greater accuracy is gained both from higher levels of polymorphism and a technique that allows finer resolution.

### 3.2 Matrilineal inheritance

It has been demonstrated that mtDNA is inherited maternally, by transmission through the egg cytoplasm (e.g. Dawid and Blackler, 1972; Hutchison et al., 1974; Hayashi et al., 1978; Avise et al., 1979; Giles et al., 1980). Present evidence suggests that there is effectively strict maternal inheritance without 'paternal leakage' (see Lansman, Avise and Huettel, 1983; Gyllensten, Wharton and Wilson, 1985), although it is still an open question (see Chapman et al., 1982; Wilson et al., 1985). Maternal inheritance was determined through cross-breeding experiments where the maternal and paternal mitochondrial genomes differed. For example, horses and donkeys have different mtDNA HaeIII restriction patterns. A cross between a female horse and a male donkey produces a mule with horse mtDNA. The reciprocal cross produces a hinnie with donkey mtDNA (Hutchison et al., 1974).

The questions of maternal inheritance and apparent haploidy of mtDNA are of central importance. There are about $10^{5}$ mitochondria in a mammalian egg, and about 50 in the midpiece of the sperm. If the sperm contributes no mitochondria to the subsequent generation, and the mtDNA in the egg is homogeneous, then mtDNA will be transmitted as a haploid genome, and only within matrilines. This would make mtDNA a powerful genetic marker for population studies. For species where there are sex biased dispersal patterns such that females tend to be philopatric (as is the case for many mammalian species: e.g. white-tailed deer - Hawkins and Klimstra, 1970; lions - Schaller, 1972; ground squirrels - Sherman and Morton, 1975), a comparison with variation in nuclear DNA would show much greater differentiation in the maternally transmitted genome (mtDNA). This could be employed to help resolve investigations where sex biased dispersal was suspected.

This interpretation depends on complete homoplasmy (no intra-individual variation in mtDNA). Theoretically heteroplasmy could arise either through mutation or by paternal contribution. There is some indication that this may occur rarely in rats (Brown and DesRosiers, 1983; Brown and Simpson 1981), cattle (Hauswirth et al., 1984), humans (Monnat and Loeb, 1985) and some other vertebrate species. This could have important consequences for the interpretation of genealogies (see discussion in Wilson et al., 1985).

### 3.3 Measures of mitochondrial genetic diversity

Within a population, isozyme variation is usually measured in terms of heterozygosity. For restriction site polymorphisms it is more appropriate to compare the average number of nucleotide differences per restriction site for two randomly chosen sequences. Nei and Li (1979) refer to this as the index of nucleotide diversity. It can be defined as

$$
\mathrm{d}=\sum \mathrm{x}_{\mathrm{i}} \mathrm{x}_{\mathrm{j}} \mathrm{~d}_{\mathrm{ij}}
$$

where $x_{i}$ is the frequency of the $i$-th sequence in the population and $\mathrm{d}_{\mathrm{ij}}$ is the number of nucleotide differences between the i -th and j -th sequences ( Nei and $\mathrm{Li}, 1979$ ). If the number of DNA segments sampled ( $n$ ) is small, then d can be estimated by multiplying by $n /(n-1)$ (Nei and Tajima, 1981).

When comparing populations, a simple measure of similarity is the proportion of fragments shared between the mtDNA digestion profiles. This is given by

$$
\mathrm{F}=2 \mathrm{n}_{\mathrm{x}} /\left(\mathrm{n}_{\mathrm{x}}+\mathrm{n}_{\mathrm{y}}\right)
$$

where $n_{x}$ and $n_{y}$ are the number of fragments for populations $x$ and $y$, and $n_{x y}$ is the number of shared fragments (Upholt, 1977; Nei and Li, 1979). Upholt (1977) has related this quantity to an estimate of the number of mtDNA base substitutions per nucleotide separating two populations ( p ):

$$
\mathrm{p}=1-\left(\left(\left(\mathrm{F}^{2}+8 \mathrm{~F}\right)^{1 / 2}-\mathrm{F}\right) / 2\right)^{1 / \mathrm{n}}
$$

where $n$ is the number of base pairs recognized by the restriction enzyme. Nei and Li (1979) derive the following estimate for this measure:

$$
d=-(\ln F) / n
$$

This is based on the assumption that $F$ can be used as an estimate of the proportion of ancestral restriction sites that
remained unchanged in both populations. These two formulations produce very similar results (see Lansman et al., 1981).

Both formulations are based on a number of assumptions:
(1) Nucleotides are randomly distributed in the mitochondrial genome. Although this does not appear to be the case (e.g. Brown. 1976). Nei and Li (1979) suggest that small deviations from randomness will not significantly alter the results. They site several examples where observed and expected values for restriction site number are in rough agreement (e.g. Kaplan and Langley, 1979; Shah and Langley, 1979).
(2) Fragment variation arises solely by base substitution. Substitutions have been estimated to be twice (Cann and Wilson, 1983) to five times (Aquadro and Greenberg, 1983) as frequent as additions and deletions in human mtDNA. Additions and deletions seem to be especially common in the region of replication (e.g. Hauswirth et al., 1984). In general, however, most variation in the mitochondrial genome seems to be attributable to base substitution (see Brown and Simpson, 1982; Greenberg et al., 1983; Avise and Lansman, 1983).
(3) Nucleotide substitution rates are the same for all nucleotides. This is clearly not the case. Most substitution seems to take place in the sequences flanking the displacement loop in the region of replication (Greenberg et al., 1983). However, Nei and Chakraborty (1976) point out that when the number of nucleotide differences per nucleotide site is small (as for intraspecific studies), this assumption does not produce any serious errors. If the distance measure is large (say more than 0.3 ), then it will be an underestimation (Nei and $\mathrm{Li}, 1979$ ).
(4) All restriction fragments can be detected, and fragments of similar length are not scored as identical. Lansman et al. (1981) suggest that enzymes which digest the genome into relatively few fragments (enzymes where $\mathrm{n}=5$ or 6 ) be used to avoid this problem. Nei and Tajima (1981) point out that the accuracy of the distance measure is enhanced by using enzymes that produce large numbers of fragments, and suggest that the results will not be greatly affected if a few small fragments are not detected.
Bottlenecks can greatly influence the level of mtDNA variability. For example, if a population of diploid animals is reduced to a single breeding pair, they will have four copies of the nuclear genome, but only one transmissable copy of the mitochondrial genome. Assuming homoplasmy for mtDNA and no paternal leakage, variation in mtDNA will be eliminated, while for brief bottlenecks significant nuclear variability can be retained (see Barton and Charlesworth, 1984; Wilson et al., 1985). This could have a dramatic affect on the interpretation of genetic distance between populations. If a rare genotype was fixed by a founder event in one of the populations being compared, the apparent distance would indicate far greater genetic division than was justified. A number of species show low levels of mtDNA variation compared to nuclear DNA, suggesting the possibility of a bottleneck period (e.g. Ferris et al., 1982; 1983; Ferris, Sage and Wilson, 1984). For example, the anomalously low level of variation in human mtDNA has led to the speculation that a transient bottleneck was involved in the formation of Homo sapiens (Brown, 1980; Johnson et al., 1983).

It has been suggested that the mean rate of divergence for the mitochondrial genome over a wide range of taxa is $2 \%$ per million years (Wilson et al., 1985; see e.g. Brown et
al., 1982; Ferris et al., 1983; Higuchi et al., 1984; Tanhauser, 1985). This estimate was derived from studies where evidence on species divergence (e.g. from fossils) was already available. Given this rate, it is possible to estimate the time of divergence between species or subpopulations, or the time elapsed since a bottleneck event. Using Nei's measure of genetic distance, Wilson et al. (1985) describe the simple relation: $\mathrm{t}=0.5 \mathrm{~d}$, where d is the mean pairwise divergence between two populations or species (in percent). Within a species, this is described as the time since two randomly picked individuals shared a common mother (see Table 4 in Wilson et al., 1985). An estimate of the long-term effective population size is given by dividing this quantity by the mean number of years per generation.

Restriction analysis of the mitochondrial genome has been used for a number of studies on variation between conspecific populations (e.g. Upholt and Dawid, 1977; Avise et al., 1979; Brown et al., 1979). A study on geographic populations of pocket gophers employed both mtDNA and standard isozyme analyses (Avise et al., 1979). Although little variation was seen at the enzyme loci, regional 'clones' were detected by digesting mtDNA with six 5 and 6-base restriction enzymes. This may reflect the matrilocal behaviour of this species. Genetic distances ( p after Upholt, 1977) were near zero within geographic areas, and tended to increase proportionally between more distant geographic populations (Avise et al., 1979).

Dizon and co-workers (Dizon, 1987) investigated mtDNA variation in the four proposed regional populations of spinner dolphins (Stenella longirostris) in the eastern tropical Pacific (after population divisions based on color morphology; e.g. Perrin, 1975). Samples were digested with six restriction enzymes and probed with ${ }^{32} \mathrm{P}$ labelled mtDNA sequences cloned from Commerson's dolphin, Cephalorhynchus commersonii (Southern et al., In Prep.). Despite variation in colour morphology, between group distance measures were not significantly greater than within group levels. The resolution of distance measures by this method is about 250,000 years assuming a rate of divergence of about $2 \%$ per million years (Wilson et al., 1985). Finer resolution is possible by applying more restriction enzymes and other methods such as radioactive end-labelling (see below).

## 4. DNA VARIATION: THE NUCLEAR GENOME

A decade of investigations into the organisation of eukaryotic nuclear genomes have revealed a variety of molecular mechanisms of DNA turnover operating in all examined species embracing the major living kingdoms. Such mechanisms both produce new types of mutation and can be involved with the dissemination of the mutations through sexual populations (see Section 1.3). In general, all mechanisms cause the gain or loss of genetic variants in the lifetime of an individual. Such small but persistent patterns of non-Mendelian segregation can affect the genetic composition of a population over long periods of time (Dover, 1982; Ohta, 1980). These mechanisms are responsible for generating variation at a significantly higher rate than variation due to point mutations, especially in non-coding regions. Consequently, methods that investigate this type of variation will offer the greatest potential for population and kinship studies.

The existence of turnover mechanisms operating throughout eukaryotic genomes, suggests that the majority of DNA is not passively accumulating point-mutations in a clock-like manner. On these grounds we do not recommend the use of DNA-DNA hybridisation techniques of so-called non-repetitive 'unique' DNA as a measure of the genetic distance between closely-related taxa (Sibley and Ahlquist, 1984). The overall flux in the genome is such that different components are diverging at different rates and only after very long periods of separation can it be expected that the gains in one region are balanced by the losses in another. Such average measurements are not of sufficient resolution for understanding the nature of variation within and between closely-related populations and species (see Dover, 1987).

### 4.1 The use of satellite DNA

Satellite DNAs are tandem arrays of repetitive sequences which can range from a few thousand to several million copies per individual, usually located at centromeric and telomeric regions of chromosomes which are heterochromatic in condensation (for reviews see Miklos, 1985; Singer, 1982; several papers in Dover and Flavell, 1982). The length of the repeating unit can vary from two to several thousand base pairs, which do not make sense from the point of view of the genetic code. They have been analysed intensely in many species, including Cetacea, because they are easily separated from the rest of the DNA by CsCl density gradient centrifugation, DNA reannealing and restriction enzyme digestion.

The evolution of satellite DNAs is generally understood, involving unequal crossing over for their de novo amplification and subsequent homogenisation with variant repeats. They show the classic pattern of concerted evolution (see above) in that species, or higher taxonomic units, can be recognised by diagnostic mutations that have spread through the DNA family. There is a great deal of controversy, however, concerning their functions. Whatever their effects may be on the structure or behaviour of chromosomes or on the expression of genes, it is unlikely that these effects can be recognised by selection because extremely large DNA families are considerably buffered from their own variant members when these are rare. Such variants need to increase in copy-number both in the family and in the population to invoke a response from selection.

Arnason and colleagues have produced an important series of papers on the satellite DNAs and karyotypes of several cetacean species (Arnason, Lutley and Sandholt, 1980; Arnason, 1982; Arnason, Purdom and Jones, 1982; Arnason and Widegren, 1984; Widegren, Arnason and Akusjarvi, 1985). These studies throw interesting light on the evolution of such families, indicating that they are clearly of different ages and that they evolve at different rates. The most recent finding concerns a 1,579 base pair repeat which comprises approximately $15 \%$ of the genome of the killer whale, making a total of $4-5 \times 10^{5}$ copies. It is also found in other delphinids. This unit is homologous at high levels of hybridisation stringencies to a 1,740 base pair repeat characteristic of all other cetacean families of species, indicating that homogenisation has occurred, sometime in the past, for a structurally shorter repetitive unit.

Restriction analysis of the 1579 satellite in other delphinids shows that the length and distribution of restriction sites have been conserved over 20-24 million
years. This is in contrast to a rapidly diverging satellite component of balaenopterid whales. Arnason and colleagues conclude from these data that the satellite DNA families are under different selection pressures, for functions that are still to be elucidated. While this might turn out to be true, we suggest that the different rates of evolution are more a reflection of the different rates of turnover in the families, coupled to differences in their size and chromosomal distribution. Studies on shared satellite DNAs amongst species of Drosophila indicate that they can evolve at different rates because it takes more time to homogenise new variants through large families spread over all chromosomes than it does for families limited to one pair of chromosomes (Strachan et al., 1982; 1985). Further, the Drosophila studies, which are based on comparisons of the precise sequences of many cloned repeats from each species, indicate that all the stages of transition can be observed during the spread of new mutations through the family, notwithstanding their apparent 'conservation'. Sequence analysis also reveals homogenised and fixed differences between closely-related species that were not observable using restriction analysis alone. In the absence of sequence data it might be premature to conclude that a satellite family is highly conserved. High similarities in overall organisation might remain between species since their last common ancestor simply because (i) unequal crossing over is operating at a rate that does not promote the homogenisation of many mutations throughout the karyotype and (ii) not sufficient time has elapsed between the species under comparison.

In conclusion, the very high copy-numbers of many satellite DNAs do not make them a useful component for studying genetic variation on a fine scale. This could be done by exhaustively cloning and sequencing representative repeats from different individuals or populations, but the time and cost would be prohibitive. Hybridisation and restriction enzyme analysis are of insufficient resolution for answering the questions posed in Section 1. As with gross DNA-DNA hybridisation techniques, we do not recommend their use.

### 4.2 Ribosomal DNA variation

(i) Sequence homogeneity and copy-number heterogeneity One of the most useful components of nuclear genomes for the identification of individuals, populations and higher units, is the multigene family coding for the 28 S and 18 S ribosomal RNAs, (for reviews see Gerbi, 1985; Arnheim, 1983; Dover, 1982; Coen, Strachan and Dover, 1982; Coen, Thoday and Dover, 1982; Coen and Dover, 1983; Fedoroff, 1979; Long and Dawid, 1980; Tautz et al., 1987). The family consists of a repetitive unit containing one copy of each of the two genes separated by an intergenic spacer, IGS, formerly called the NTS (Fig. 1). The compound unit of genes and spacer can be repeated from several hundred to thousands of times in tandem arrays that can be on several pairs of chromosomes, depending on species.

The utility of the rDNA gene-family for the study of natural variation, population structure and breeding behaviour derives from two features.
(i) Each spacer is further divided into a tandem array of subrepeats whose precise lengths differ between species. Some spacers, such as those in Drosophila and Xenopus, contain several different arrays of subrepetition; that is units of different length and sequence can be repeated within the spacer. The organisation of the rDNA unit in Drosophila is depicted in Fig. 1.


Fig. 1. Top. An rDNA unit of $D$. melanogaster, with two genes (18S and 28S), an internal spacer (ITS) and the main intergenic spacer IGS. Some units have an insertion (INS). The species-specific Alu I sites (Alu I is a restriction enzyme that recognizes the 4 base sequence AGCT) occur in each of a 240 base pair subrepeat: other closely-related species have the same subrepeat array but with other diagnostic restriction sites that have been homogenised throughout the subrepeat array, all rDNA units and all individuals, (see text). Below. An expanded version of the IGS of D. melanogaster showing three arrays of subrepeats in this species. Other closely-related species do not have the 90 base pair repetition because of the operation of slippage generating a high level of 'cryptic simplicity' (scrambled short DNA motifs) - see Section 1.3 (iii).
(ii) Unequal crossing over is known to be occurring at both levels of repetition (Fig. 2). Unequal crossing over at the periodicity of the 240 base pair subrepetition in the spacer leads to continual gain-and-loss of numbers of subrepeats, which can be detected by restriction enzyme sites that lie outside the array of subrepeats and between which the length of DNA is longer or shorter (spacer length heterogeneity). Unequal crossing over at the periodicity of the whole rDNA unit leads to variation in the copy-number of a particular length generated at the lower level. This can be detected by the intensity of bands in a gel restriction digest of the rDNA family. Hence, the position of a band is indicative of the length of the fragment (which is a reflection of the number of spacer subrepeats); and the intensity of a band is indicative of the number of whole units of a given length.


Fig. 2. Top. Unequal crossing over at the periodicity of the subrepetition within each spacer (see Fig. 1) generates variation in copy-number of subrepeats reflected as longer and shorter distances between restriction sites on each side of the array.
Below. Unequal crossing over at the periodicity of the whole rDNA unit (see Fig. 1) generates variation in copy-number of the different spacer lengths generated by the first level of unequal crossing over. This is reflected in the intensities of bands of different lengths in a hybridisation gel.

Measured rates of unequal crossing over both within and between different chromosomal arrays are approximately $10^{-4}$ per generation per rDNA unit in Drosophila (Coen et al., 1982a, b; Coen and Dover, 1983) and yeast (Szostak and $\mathrm{Wu}, 1980$ ). These rates are faster than the base substitution rate (and hence lead to the homogenisation of mutations through the arrays and species - see Section 1.3); but are too slow to generate new spacer lengths at a rate which would disturb their use for the identification of parent-offspring. This disparity in rates is true also for the hypervariable minisatellites used for DNA 'fingerprinting' and is the basis for high probabilities of correct identification (Section 4.3).

In a study of rDNA variation amongst species of Drosophila, Coen and colleagues showed that the 500 copies per individual were homogeneous for mutations that were diagnostic for a species (Coen et al., $1982 \mathrm{a}, \mathrm{b}$ ). For example, each of the spacer subrepeats in each of the spacers in $D$. melanogaster contained a restriction enzyme site (Alu 1 - see Fig. 1) not present in D. simulans and several other related sibling species. In humans and mice it has been shown that the rDNA families are divisible into subfamilies on the basis of diagnostic restriction sites (Arnheim, 1983). These might represent partially homogenised mutations or they might reflect the restriction of unequal crossing over to a subset of the available repeats. Chromosomes, for example, might be a natural barrier to unequal crossing over in species where the rDNA array is divided amongst several non-homologous pairs of chromosomes. In humans, however, the subfamilies are evenly distributed amongst the five chromosomal locations indicating that in this case the chromosome is not a barrier to exchanges by unequal crossing over, or possibly gene conversion. The opposite situation has been found in mice rDNA in which subfamilies are chromosome specific, although mouse satellite DNA has subfamilies shared amongst all chromosomes (Brown and Dover, 1981). Such studies indicate that the evolutionary history and subsequent distribution of genetic variation from chromosomes upwards, is different for each family in each species; no generalisations can be made.

Despite the extensive homogenisation (or partial homogenisation) of mutations in rDNA families and the corresponding reduction of variation within but not between species, there is a great deal of variation to be exposed due to the gain-and-loss of copy-numbers of spacer subrepeats and whole rDNA units, as described above. This variation can be detected with the use of appropriate restriction enzymes on whole DNA and probing the resultant gels with different regions of the rDNA unit by the Southern hybridisation technique. Using such techniques it was shown that individual X and Y chromosomes in $D$. melanogaster (which carry the rDNA arrays) could be uniquely identified, (Coen et al., 1982 a , b). Further, the molecular characterisation of rDNA length and copy-number variation during an experiment involved with the selection of high and low bristle number was able to show that the response to selection was based, amongst other things, on variation in the copy-number of rDNA, due to the activities of unequal crossing over during the period of selection (Coen and Dover, 1983). These experiments were the first to be able to positively identify individuals using DNA fingerprints based on rDNA variation. Further, they showed for the first time that very small amounts of tissue (single flies) could be successfully
monitored at the molecular level using appropriate DNA miniaturisation techniques.

## (ii) Copy-number heterogeneity and population identification

Over the past few years rDNA has been extensively used as a genetic marker in diverse species, in particular in plants. These are reviewed in two recent reports by Flavell et al. (1986) and Learn and Schaal (1987) analysing populations of tetraploid wild wheat (Triticum dicoccoides) and Clematis fremontii, respectively. These two studies raise further interesting questions concerning the role of natural selection and genetic drift in the interaction with molecular drive during the distribution of rDNA variants in the populations.

For example, Flavell et al. (1986) have examined the distribution of spacer length variants in 112 plants taken from 12 populations for which allozymic variation encoded by 50 gene loci had been previously established. Populations of wild wheat are geographically structured and are predictable by ecological and allozyme markers (Nevo, 1983; Nevo et al., 1982). The distribution of allozymes suggested that selection was responsible for some of the differences in localities with different climates and soil types.

The rDNA family is distributed on two pairs of chromosomes, and spacer length variation is due to the gain-and-loss of a 135 base pair subrepetition within the spacer. The results of the survey show that natural populations of $T$. dicoccoides display a wide spectrum of spacer lengths, with some populations being very homogeneous (that is, surprisingly all arrays of rDNA on the two non-homologous chromosomes have the same spacer length); whilst other populations have either intermediate levels or very high levels of heterogeneity. The allozymic and rDNA diversities are significantly intercorrelated both between themselves and with the climatic variables.

The highly heterogeneous populations had nine or more different lengths of spacers, whereas the homogeneous populations displayed a single length that was the most frequent length in the heterogeneous populations. These results emerged with the use of a single restriction enzyme Taq 1. The further use of ECoRI + Bam HI, Dde I, and Hinf I showed that there are least 3 major types of rDNA repeat in the homogeneous population. This illustrates how the use of one restriction enzyme underestimates the heterogeneity within the individual. However, the use of Taq I has not underestimated the heterogeneity within the population, since, with all enzymes used, all individuals within a homogeneous population were identical. The additional enzymes were revealing additions and deletions of DNA within the spacers, but outside the array of 135 base pair subrepeats.

The average number of electrophoretic alleles per population, the proportion of polymorphic loci per population and levels of genetic diversity (using the measures derived by Nei, 1975; see Section 2) are highly correlated with rDNA variables, using Pearson correlations, in particular between the genetic diversity index and the number of independently occurring spacer lengths.

A full explanation of the forces responsible for these distributions is not possible from the data available so far. T. dicoccoides displays a highly subdivided population structure with only limited gene flow between
semi-isolated populations, which can be characterised by peaks of locally common alleles. Populations with high allelic variability and rDNA variants are situated in climatically highly fluctuating regions.

Is selection acting on the rDNA length variants directly or on alleles tightly linked to them? Species of Drosophila, Xenopus and wheat are known to contain functionally important signals for transcription within the spacer subrepeats, the copy-number of which affects transcription levels (Moss, Mitchelson and de Winter, 1985; Reeder, 1984; Dover and Flavell, 1984). It could be that this is the functional basis on which selection can act. However, it is unlikely, as stressed in previous sections, that selection can act on the first variant spacer occurring in a family of several hundred members. Some appreciable level of homogenisation and fixation would need to take place, as a consequence of unequal crossing over in this case, before appropriate effects on phenotype are 'visible' (Dover, 1982; 1986a; Ohta and Dover, 1984; Ohta, 1980; 1983). Some means of generating the same spacer length variant on different chromosomes must also be operating. From what is currently known in $D$. melanogaster this, too, could be unequal crossing over.

The finding that the homogenised spacer length in the monomorphic populations is the one most frequent in the polymorphic populations suggests too that an element of chance fixation in small populations has also been involved by a combination of genetic drift and molecular drive. This is in essence the conclusion drawn by Learn and Schaal (1987) on the population subdivision of length variants in Clematis fremontii.

Clematis fremontii is a long-lived perennial plant established in prairie and woodland glade populations in Missouri. Prairie and glade populations have been given subspecific status. Glades are prairie remnants isolated about 5,000 years ago, forming islands in the surrounding forest. Analysis of leaf shape suggests that glade populations are genetically subdivided. Both cross and self-fertilisation takes place although the amount of selfing is not known.

Along a hillside transect of 16 adjacent $10 \times 10 \mathrm{~m}$ quadrats in a given glade, 217 plants have been sampled (Learn and Schaal, 1987). The numbers and lengths of spacer variants were examined in at least 12 individuals from each quadrat. Scanning densitometry of autoradiograms has been used to estimate the relative proportions of each length variant. Individuals contain several repeat lengths with an average of 2.67 per individual. This is at least two times smaller than the number of length variants found in Drosophila. The most frequent length variants amongst the quadrats were also the same length variants with the highest copy-number within individuals. However, a separate class of length variants could be detected which were relatively rare or completely absent in several quadrats but which can reach up to $25 \%$ of the repeats within some individuals in other quadrats. The number of plants with high copy-numbers of an infrequent variant at the population level can range from a few to $50 \%$ of the plants in a given quadrat. Other relatively rare variants do not quite reach $10 \%$ of the repeats within an individual. Results from statistical analysis (nonparametric Kruskal-Wallis test) show a significant subdivision of rare variants amongst quadrats.

Since rDNA repeats constitute a heterogeneous collection even within a single individual, genetic diversity can be apportioned into within-individual, within-quadrat
and within-population components using Shannon's information measure (Lewontin, 1972; see Section 2.3). Using this measure, Learn and Schaal (1987) show that an average individual contains $65.6 \%$ of the total population diversity; and that the average quadrat displays more than $95 \%$ of the total diversity. These two numbers are related in that the degree of length variability generated by unequal crossing over and other deletion/expansion events within an individual would affect the overall population heterogeneity if a few variable individuals are involved with the establishment of isolated populations of the $C$. fremontii glade subspecies. The existence of unusual length variants in some parts of a population but not others, and at relatively high copy-numbers per individual, indicate that population differentiation is due to genetic drift and restricted gene flow. Hence, the final distribution of length variants is an outcome of molecular drive and genetic drift in this instance. The rDNA length variant patterns parallel the patterns seen for morphological and allozyme markers in other populations of this species.

In conclusion, it is clear from studies in Drosophila, Triticum and Clematis that the rDNA multigene family can provide genetic variability that can usefully identify taxa from individuals to species and reveal the history of genomic and ecological events that have shaped well-studied populations. High levels of length and copy-number variability, which are the products of two nested levels of unequal crossing over (see Fig. 2; and Coen et al., 1982a, b) are useful for monitoring population changes at a microlevel; whereas the low levels of sequence variability, due to the homogenising consequences of unequal crossing over provide an unambiguous means of identification of genetically distinct higher taxa such as species.

### 4.3 Hypervariable minisatellites and DNA fingerprinting

 (i) Minisatellites and restriction fragment length polymorphisms, (RFLPs)Considerable and justifiable excitement has arisen over the discovery of hypervariable minisatellite regions in human DNA that can be used to identify individuals unambiguously, by the so-called DNA 'fingerprinting' technique (Jeffreys, Wilson and Thein, 1985). The term fingerprint does not do justice to the power of the technique which, in addition to identifying individuals from very small amounts of bodily fluids, can also positively identify the closest relatives of an individual. All other available genetic markers can only eliminate individuals from being close relatives; they cannot positively identify individuals and their genetic relatedness quickly and efficiently with the same verity as the available minisatellite DNA probes. Studies on the application of this technology to the description and differentiation of populations are underway but as yet incomplete. However, as described below, the analysis of breeding systems is greatly facilitated by utilising DNA fingerprints to establish pedigrees, and this is an important component to understanding and predicting dispersal patterns and other mechanisms of population mixing.

Minisatellites derive their name from their existence as relatively short tandem arrays of repeats scattered on all but the sex chromosomes. Hypervariability is detectable as variation in copy-number of repeats at the different loci. This is the same as the high variability in copy-number of subrepeats in each rDNA unit which gives rise to spacer
length variation, (see Section 4.2 and Fig. 2). As with the rDNA, differences between individuals with respect to the repeat copy-number in each array can be detected with the use of a restriction enzyme that cuts outside of the array and not within any of the repeats. Hence, variation in the distance between the two given restriction sites gives rise to DNA fragments in a gel which can be detected after hybridisation to a probe of the repetitive unit.

Before the advent of minisatellite probes, the detection of nuclear DNA variation relied almost exclusively on the availability of polymorphisms at the target sites of restriction enzymes. RFLPs (restriction fragment length polymorphisms) have been used extensively for several genetic purposes, but their resolving power, except in the case of the more rapidly diverging mitochondrial genomes, (see Section 3), is weak. RFLPs usually arise by point-mutations whose rate of production is approximately two orders of magnitude slower than the rate of variation in copy-numbers of repeats generated by unequal crossing over or slippage (see Section 1.3). As the mean heterozygosity of the human DNA is low (approximately 0.001 per base pair-Jeffreys, 1979; Ewens, Spielman and Harris, 1981), few if any restriction enzymes will detect a RFLP at a given locus. Even when detected, most RFLPs are only dimorphic with a heterozygosity which can never exceed $50 \%$, and which is usually much less. As Jeffreys and colleagues point out (1985a) all such RFLPs will be uninformative in pedigree analysis whenever critical individuals are homozygous.

Some improvement of the RFLP approach is possible if probes can be isolated from libraries of genomic clones which by chance cover polymorphic regions, or with a carefully chosen set of tightly linked RFLPs producing chromosomal haplotypes (Smouse and Chakraborty, 1986). This former type of improvement has been used recently for the analysis of multiple paternity and maternity in single broods of the lesser snow goose, (Quinn et al., 1987). Conventional genetic markers such as allozymes are usually inadequate for such determinations, especially in avian species. The isolation of 17 probes from genomic libraries of lesser snow geese that identify RFLPs has allowed for some correct parental identifications. However, notwithstanding such improvements, the degree of available variability is low relative to that revealed by minisatellite copy-number variation, a difference which bears significantly on the resolving power and statistical reliability of genetic identification of the two techniques*. Considering the poor resolving power of RFLP detection methods, compared to the ready availability of both naturally occurring and synthetic probes for the detection of several different minisatellites, we suggest that minisatellites be investigated in preference to RFLP analysis, especially for pedigree studies.

The first human minisatellite was isolated by Jeffreys and his coworkers, which comprised four tandem repeats of a 33 base pair sequence in the intron of the myoglobin gene. The four repeats were flanked by a 9 base pair direct repeat indicating that it had moved into this locus from elsewhere [see Section 1.3 (i)]. A pure repeat probe was prepared by purification of a single 33 base pair repeat element followed by head-to-tail ligation and cloning of the resulting polymer of 23 repeats. Using this polymer as a probe against Hinf I or Hae III restricted total DNA,

* For a full discussion of the statistical problems associated with paternity identification using linked and unlinked RFLPs see Smouse and Chakraborty (1986).
multiple DNA fragments were detected as well as the parent DNA fragment from the myoglobin gene. Differences in the size of fragments both within and between individuals are a result of copy-number variation of the repeats at different loci, because neither Hinf I nor Hae III cut within the repeats themselves. Fragments are transmitted in a Mendelian fashion from parents to progeny in that each band in a progeny can be assigned to one or other parent.

Cloning and isolation of the minisatellite from different loci shows that amongst the eight randomly chosen clones, the arrays differed in length from 3 to 29 tandem copies of a repeat whose length ranged from 16 base pairs to 64 base pairs. Repeats from different loci are related by a shared central 'core' sequence, GGGCAGGAXG, which is similar in length and in GC content to the chi-sequence, a signal for generalised recombination in E. coli. This 'core' could be acting as a recombination 'hotspot' during unequal crossing over or slippage, so generating high variability in repeat copy-number. The presence of different sequences flanking the core and some variation in the core itself permits the use of probes that can detect different patterns of minisatellites, substantially increasing the resolving power of the technique and reliability of identification. Each probe hybridises only to arrays of repeats of its own constitution.

## (ii) Assessment of genetic relatedness and population heterozygosity

The use of two different repeat probes ( 33.15 and 33.6, see Jeffreys et al., 1986) in a large sibship affected by neurofibromatosis and a more extensive pedigree segregating for two different haemoglobinopathies, reveals up to 41 different heterozygous DNA fragments from each parent. Most fragments could not be paired as alleles, to an extent which suggests that the resultant DNA fingerprints are together derived from approximately 60 heterozygous loci which is equivalent to approximately 120 variable fragments, only a proportion of which can be scored in a given individual. Excluding a few allelic and linked DNA fragments, up to 34 unlinked loci have been examined simultaneously. These are scattered over most or all of the autosomes.

Table 3
Range of fragments found using two repeat probes (see text). The data are taken from Jeffreys et al. (1986, Am. J. Human Genet. 39: 11-24)

|  | Father |  |  |  |  |
| :--- | ---: | ---: | ---: | ---: | ---: |
| Probe | 33.6 | 33.15 |  | Mother |  |
| No. fragnents scored ( n ) | 24 | 17 |  | 16 | 16 |
| No. allelic pairs (a) | 3 | 3 |  | 2 | 4 |
| No. linked pairs (b) | 1 | 0 |  | 1 | 0 |
| No. different loci scored (L) | 20 | 14 |  | 13 | 12 |
| Estimated total no. loci (N) | 43 | 23 |  | 27 | 16 |

Table 3 contains the results* of the range of fragments found in a neurofibromatosis family of a father, five sons and six daughters using probes 33.6 and 33.15. These data are presented here as an illustration of the resolving power of the technique and the formulae developed by Dr. John F.Y. Brookfield for data analysis.

[^8]Maternally derived fragments can be readily identified as fragments which are present in some of the 11 siblings but which are absent from the father. Paternal fragments can similarly be identified. Using both probes, it was possible to score the segregation of 41 paternal and 32 maternal fragments (Table 1).
The number of different loci $(\mathrm{L})$ scored is given by n-a-b. The entire DNA fingerprint, including unresolved and therefore unscored fragments, is derived from N heterozygous loci ( 2 N fragments). Assuming that the (n-b) distinct fragments scored are a random sample of the 2 N bands in a DNA fingerprint, then the estimated total number of hypervariable loci N detected by a given probe is related to the number of allelic pairs a by

$$
\mathrm{N}=\frac{1}{2}\left[\frac{(\mathrm{n}-\mathrm{b})(\mathrm{n}-\mathrm{b}-1)}{2 \mathrm{a}}+1\right]
$$

Identification of allelic and linked pairs of fragments is possible by pairwise comparisons of the segregation patterns of all paternal or maternal fragments amongst the 11 children. Several instances of allelic pairs of both paternal and maternal fragments can be identified by careful screening of the gels. By eliminating all such alleles and linked fragments, Jeffreys and coworkers conclude that 34 and 25 distinct loci have been scored in the father and mother respectively. For approximately $80 \%$ of the loci, only one of the two alleles is resolved, and the second allele is probably located in the poorly resolved complex of shorter fragments at the bottom of the gel. From the proportion of bands that can be paired into alleles, it is possible to estimate that the total number of heterozygous loci present in the entire DNA fingerprints detected by probes 33.6 and 33.15 is approximately 43-66, of which roughly half can be scored in each parent (Table 1). It is not possible to determine allelism between paternal and maternal fragments in this sibship.

From the results described above it is clear that the correct parentage of individuals can be deduced, even if one of the parents, usually the father, is not known. This provides a powerful way for unravelling complex breeding behaviours in cetacean species. However, we need to turn again to the human situation in order to illustrate the statistical robustness of the DNA fingerprint technique. The case involves the positive identification of a boy's father from fingerprints derived from the boy, his mother and three other progeny of the mother all of whom had the same father. The details are given in Jeffreys, Brookfield and Semenoff (1985a). Fingerprints were obtained with the use of the two probes described earlier.

The father's DNA fingerprint could be reconstructed from paternal-specific fragments present in at least one of the three undisputed siblings, but absent from the mother. Of the 39 paternal fragments identified, approximately half were present in the boy. Since DNA fragments are seldom shared between unrelated individuals, this strongly suggests that the boy has the same father. After subtracting these paternal-specific fragments, there remained 40 fragments in the boy, all of which were present in his mother.

The probability of a correct identification can be calculated as follows. It can be shown from the distribution of bands of unrelated individuals that the mean probability that a fragment in a DNA fingerprint of one person is present in a second individual, selected at random, is 0.2 . The corresponding estimate for the father and mother is 0.26 (see below). The boy's fingerprint contains 61
fragments, all of which are present in mother and/or father. If the boy is unrelated, the probability ( $x$ ) that each of his bands is present in these parents is $1-(1-0.26)^{2}=0.45$. Hence the probability that mother and/or father by chance possess all 61 of the boy's bands is $0.45^{61}=7 \times 10^{-22}$. Similarly, the 25 fragments in the boy that can be unambiguously assigned to the mother, have only a $2 \times 10^{-15}\left(0.26^{25}\right)$ chance of coincidental occurrence (Jeffreys et al., 1985a).
The quantification of DNA fingerprints is important, and the following illustrates the type of analysis that can be made (as presented in Jeffreys et al., 1985a). A total of 61 fragments are scored in the mother, compared with 39 fragments known to be inherited from the father. One-eighth of the father's heterozygous DNA fragments will not be transmitted to the three siblings, and thus the corrected estimate for the number of parent-specific fragments is $39 \times 8 / 7=45$. Since the total number of fragments in mother and father should be approximately equal, the number of fragments in mother shared by the father is approximately $(61-45)=16$. The mean probability of band sharing ( x ) in mother and father is $16 / 61$ $=0.26$.
Given the range in x from 0.2 to 0.26 , then the probabilities of sharing all bands can be calculated as follows. Since almost all fragments are inherited independently, the maximum probability that all $n$ fragments in an individual are present in a second random individual is $\mathrm{x}^{\mathrm{n}}$. With regards band sharing in sibs the following applies. If shared bands always represent identical alleles at the same locus, then, assuming that all alleles have equal frequencies, $x$ is related to the allele frequency $q$ by $x=2 q-q^{2}$. At Hardy-Weinberg equilibrium, the probability that a band in an individual is also present in a sib can be shown to be

$$
\frac{\left(4+5 q-6 q^{2}+q^{3}\right)}{4(2-q)}
$$

The other extreme case is that bands shared by unrelated individuals are never allelic (that is, that there are many loci at which a band with a given electrophoretic mobility may be found). Then q can be defined as the probability that a given band will be found in a random gamete from the population. As before, $x=2 q-q^{2}$, but the probability of band sharing between sibs can now be shown to be

$$
\frac{\left(1+q-q^{2}\right)}{2-q}
$$

For $x=0.26, q$ is 0.14 and the probability of band sharing is 0.62 in the first case (alleles) and 0.60 in the second case (non-alleles). (These calculations are given by Jeffreys et al., 1985a).
Lynch (1988) cautions that various problems in addition to the co-migration of non-allelic markers, could produce high variances when similarity between DNA fingerprint patterns is used for kinship estimation. These include having a finite number of alleles, possible homosygosity at some loci, loss of resolution for low molecular weight markers and possible linkage.
It is important to stress at this point that the application of traditional analyses of allele frequencies and probabilities of sharing arises from the observation that the DNA fingerprint fragments are, to all intents and purposes, stably inherited in a Mendelian manner. As described above in Sections 1.3 and 4.2 this holds true so long as the rate of production of new length variants is as
low as $0.5-1.5 \times 10^{-4}$ per gamete per kilobase of minisatellite. At this low rate, it is highly improbable that the DNA fingerprints of a few related individuals in a defined pedigree over a few generations will reveal the accumulation of novel fragments via the non-Mendelian transmission consequences of unequal crossing over and/or slippage. After long periods of time, however, the multiple length variants produced by either mechanism would have accumulated in the population, which is why the minisatellites are currently observed to be hypervariable. It might not be entirely justified, therefore, to employ methods of analysis of population variation based on Hardy-Weinberg equilibria, for this supposes that the population is essentially 'Mendelian', whereas in truth the frequencies of variants in the population are determined by both the Mendelian transmission of chromosomes and the non-Mendelian behaviour of the DNA. Fortunately, in practice this should not overly influence the interpretation of results.

In general, hypervariable minisatellites are proving to be a powerful tool for understanding the genetic relationship between individuals and, as such, will aid in the resolution of certain problems associated with cetacean species (see Section 4.3.v). Their use for understanding the genetic differentiation of populations is more problematical when used in isolation of other genetic markers, and is still in its infancy. More empirical data are required to know whether separate populations can be uniquely identified by an 'average' consensus DNA fingerprint. The distribution of band frequencies in a population of some species is roughly polynomial with most bands being rare and a few being more common (especially low molecular weight bands) (Amos and Hoelzel, in prep.). It remains to be seen if these less variant bands can be used to distinguish populations. This would best serve as a population 'marker', however, as the interpretation of genetic distance and divergence times requires a better understanding of the process generating variation than is currently available for mini-satellite sequences.

## (iii) The use of single-locus hypervariable probes

For the specific purposes of linking minisatellite loci to other genetic markers, or for the identification of races and populations, use can be made of clones of locus-specific arrays of minisatellites. One particular array has proven useful in this respect because it is an extremely polymorphic locus (heterozygosity $=97 \%$ ) isolated from a single band (locus) in a human DNA fingerprint (Wong et al., 1986). The locus shows extreme length variation due to allelic variation in the number and slight differences in the length of the repeat unit. In a random sample of 158 chromosomes, one common and 76 rare alleles could be resolved. The estimated rate of production of new length variants (assuming Ne for humans is approximately $10^{4}$ ) is 0.002 per gamete. The average length of minisatellite DNA at the locus is 5 kilobases, and thus the rate per kilobase is $4 \times 10^{-4}$ which compares to rates of $10^{-4}$ from other loci (see above). The increase in fragment number from this locus relative to other loci decreases the probability that two unrelated individuals will share the same fragment.

Too much variability can be detrimental for race and population identification, because the high rates of generation of new variants in each population will tend to obliterate any population-specific variants that might have differentiated the populations at the time of inception. This
would be true particularly in the case where the periodicity of unequal crossing over can vary greatly from one to many repeat units (see Section 4.2), hence generating a high range of length variants (copy-number variation) in each array. The use of probes taken from loci that show lower levels of polymorphism would be more appropriate for the distinction of higher taxonomic units. These could be used in conjunction with rDNA and RFLPs in order to generate sets of data that overlap and cover taxa from individuals to species.
(iv) Locating other minisatellites with bacteriophage M13 and synthetic probes
The discovery of one hypervariable minisatellite by Jeffreys and co-workers has led to the identification of other minisatellites using a range of probes. These have included naturally occurring probes isolated from specific cloned regions of a genome; the surprising use of bacteriophage M13, and synthetic simple sequence DNA [see Section 1.3 (iii)].

Weatherall and colleagues at Oxford (Jarman et al., 1986) have characterised a highly polymorphic DNA region 8 kilobases downstream of the human alpha globin gene complex. It is composed of an array of 17 base pair repeats, the number of which varies from $70-450$ from allele to allele. The sequence itself is highly conserved within and between alleles. Furthermore, the sequence itself identifies a core oligonucleotide GNGGGNACAG (where $\mathrm{N}=$ any nucleotide), that is common to three previously characterised hypervariable regions. At reduced hybridisation stringencies a probe of the hypervariable region at the gamma-globin complex detects multiple Mendelian inherited DNA segments, suggesting that it too may represent one member of a dispersed minisatellite family. There is no homology, however, between this family and that isolated by Jeffreys.

Similarly, no homology has been found between the original minisatellite DNA family and several others that have been discovered with the use of synthetic probes of oligonucleotide consensus sequences from arrays of repeats near known genes within the human genome (Nakamura et al., 1987). In addition, 16-base and 20-base oligonucleotides have been synthesised that correspond to a portion of the X -gene region of hepatitis B virus, on the basis of an apparent similarity of these sequences to the consensus sequence of the myoglobin family. A total of six synthetic probes revealed 1,000 clones per genome in human genomic cosmid libraries with each probe hybridising to its own set. Selected clones were then used as hybridisation probes using the Southern blot technique against individual genomes, revealing high levels of copy-number variability at many genomic loci. Further subcloning of defined regions of the original clones could produce probes that hybridised to single loci only. Of 372 clones tested so far, $77(21 \%)$ have revealed hypervariability at specific loci, with nearly $90 \%$ of these showing three or more alleles. Heterozygosity is over 70\% at loci with three or more alleles.

A recent report by Vassart et al. (1987) describes the findings of hypervariable minisatellites in human and other animals with the wild-type M13 bacteriophage, provided no competitor DNA is used during the hybridisation. The effective sequence in M13 was traced to two arrays of 15 base pair repeats, [corresponding to (Glu, Gly, Gly, Ser) $n$,] within protein III gene of the bacteriophage. Nine
unrelated individuals displayed different patterns with approximately 30 bands in each, whilst monozygotic twins were indistinguishable. The patterns were clearly different from those obtained with Jeffreys' probe. Highly polymorphic patterns were also obtained with DNA of bovine, equine, murine and canine origin. These results lead to the surprising finding that one of the most commonly used DNA vectors is able, by itself, to detect hypervariability in a variety of mammals.

Synthetic probes of minisatellites have been successfully made and used starting from 14 base pair oligonucleotides with subsequent concatemerisation into tandem arrays (Vergnaud, 1987). Three different repeat sequences were synthesised: the first is a random polypurine tract, and the second and third are based on a triplet GCA or ATT, with two and four variations, respectively. All three probes are simple sequence in composition. They were chosen because of the existence of extensive cryptic simplicity throughout eukaryote genomes, (Tautz, Trick and Dover, 1986). Unrelated individuals produced polymorphic bands when probed with the synthetic constructs. Although the number of bands was low, there was sufficient variation to follow a simple pedigree. It should now be possible to synthesise a wide range of probes for the detection of hypervariable minisatellites.

All the above three studies indicate that most regions of eukaryote genomes are composed of runs of short repetitive motifs of many different sequence compositions. The molecular techniques are now available to uncover this variation in a systematic manner and to harness it to answer important questions of the breeding, social and migratory behaviour of mammalian species. To date only the tip of the iceberg has been revealed.
(v) DNA fingerprints in seals and whales: future potential Recent attempts to DNA fingerprint pilot whale, killer whale and grey seal have proved successful (Amos, Hoelzel and Dover, unpublished data; Hoelzel and Amos, 1988). There are no major technical problems for a large scale survey of genetic variation in natural populations of sea mammals. To date, the use of the two human minisatellite probes described above detect up to 50 polymorphic bands unique to each probe in grey seals and $40-50$ in pilot and killer whales. In the former species the bands have been shown to be independently inherited in a Mendelian manner, permitting the identification of parent-offspring relationships, (Amos, Anderson and Dover, in prep.). Small samples of tissue collected from whales can be adequately preserved in salt buffers for up to several weeks in mild refrigeration before DNA extraction. The success of human probes against whales and seals indicates that it can be expected with confidence that much more genetic variation is to be uncovered with the use of whale-specific or seal-specific probes isolated from cloned genomic libraries. Curiously, the human probe 33.6 hybridises to the 1579 satellite described by Arnason (see Section 4.1). This makes the establishment of whale-specific probes more urgent. The same urgency is true for rDNA probes. Currently available clones of the highly conserved 18 S and 28 S genes from Drosophila can be used to locate the rDNA arrays in sea mammals. Earlier work has located the rDNA genes of cetacean species in certain genomic fractions isolated in cesium chloride gradients (Arnason et al., 1977). It will then be a relatively easy matter in the appropriate molecular laboratory to
proceed with the isolation and characterisation of whale rDNA units. In keeping with all other examined species so far, it is to be expected that species-specific sequence homogeneity patterns will characterise the Cetacea, with individual and population identities established by high variability in rDNA spacer lengths (see Section 4.2).

The correct interpretation of variation in hypervariable minisatellites and rDNA spacers for cetacean species will be facilitated by some knowledge of the breeding behaviour of the species in question. This is especially important when genetic variants are to be used to statistically distinguish populations. For populations where complete information on the association of individual animals is available (e.g. through long-term individual-recognition studies), and known cow-calf pairs can be identified, this can be determined by using DNA fingerprinting to identify paternity and extending genealogies. For example, through the analysis of DNA variation in the population of killer whales in Puget Sound (Washington, USA) where matrilineal genealogies are known for up to three generations, and within three apparently isolated populations a finite set of potential fathers can be identified (see Bigg, 1982; Osborne, Felleman and Heimlich-Boran, 1985; Hoelzel, study in progress). From such surveys it will be possible to identify the number of allelic and non-allelic fragments produced by given minisatellites to a good approximation, using the analytical procedures devised from the human studies (see Wong et al., 1986). On this basis, the frequencies of alleles at each locus can be assessed and their closeness of fit to a Hardy-Weinberg equilibrium distribution can be ascertained. In addition, overall pod and population heterozygosities can be measured. The data can then be analysed further for a variety of purposes using methods described in full in Section 2.3-9. Genetic markers from rDNA spacer repeat sequences or mtDNA haplotypes will facilitate the differentiation of populations without specific knowledge of the Mendelian segregation behavior of the markers (see above).

It can be expected that comparisons of all such parameters between individuals, between pods and between stocks (populations) will yield information on, for example the genetic relatedness of individuals within a pod; the number of shared fathers; family sizes; the reproductive success of fathers; the gene-flow between pods; and differences between populations in the types and frequencies of loci and alleles for each minisatellite DNA family.

In general, the hypervariable minisatellites are most applicable for resolving genetic relatedness between individuals in a breeding group. Further, they could be employed as individual markers for mark-recapture census studies, as each individual will have a 'DNA fingerprint' that is very unlikely to be shared with any other whale in the population (see above). Their use can be extended upwards to the population level, [see last paragraph Section 4.3 (ii)] although other markers more amenable to the identification of isolated and semi-isolated populations can be obtained with the exploitation of rDNA sequence and length variability, along the lines employed in populations of $T$. dicoccoides and C. fremontii [see Section 4.2 (ii)]. Interestingly, rDNA length variation can be extended downwards to the individual and chromosome level, and hence the overlap between the two data-sets derived from the minisatellites and rDNA makes them a powerful combination.

## 5. PRACTICAL CONSIDERATIONS

### 5.1 Sample collection

The requisite conditions for collection and storage of material varies considerably for the different techniques. Isozyme studies require small to medium sized samples ( $<1-5 \mathrm{~g}$ ) depending on the tissue. These should be frozen at $-20^{\circ} \mathrm{C}$ soon after collection. If the material is to be stored for longer than a few weeks prior to analysis, it should be kept in liquid nitrogen or in an ultracold freezer $\left(-70^{\circ} \mathrm{C}\right)$. Samples can be stored this way without any apparent degradation of enzyme activity for several months. Unnecessary thawing and refreezing should be avoided. For blood samples, plasma and red blood cells should be separated by centrifugation before freezing and stored separately. As little as $50-100 \mu \mathrm{l}$ is sufficient once an analytical protocol has been established.

Captive whales can be readily sampled for blood, as blood samples are routinely collected from these animals to monitor their health. Stranded animals are often subject to rapid degradation due to overheating. If the stranding is very recent, or if the ambient temperature is near freezing or below, samples should be collected from various organs (especially liver, heart and muscle; this also applies to incidental takes). Otherwise skin samples should be collected, as this material does not degrade as quickly as the internal tissues. For isozyme analysis it is useful to obtain large samples from various tissues to establish enzyme and buffer systems for a particular species and tissue. One of us (ARH) has conducted enzyme electrophoresis on skin samples from several cetacean species (Globicephala macrorhynchus, Physeter macrocephalus, Delphinus delphis, Balaenoptera physalus). Globicephala and Physeter samples were from stranded animals. Twenty six putative enzyme loci were investigated, of which 24 were evident in dermal samples.

For surveys of wild populations, it is possible to collect small biopsy samples in the field without restraining or harming the subject animal. This is done by firing a dart with a rifle (e.g. Winn, Bischoff and Tarushi, 1973), speargun (e.g. Aguilar and Nadal, 1984) or bow (e.g. Dorsey, Hoelzel and Stern, 1982). The dart is typically a metal cylinder about 6 mm in diameter and 25 mm deep with a base-plate and a screw mount for attachment to a firing shaft. The leading edge is sharpened, and internal barbs (Dorsey et al., 1982) or a 'butterfly' valve (Aguilar and Nadal, 1984) retain the sample as the dart is withdrawn. One of us (ARH) has used this technique with minke whales and killer whales. The response to impact is minimal. With a bow, close range is required $(10-20 \mathrm{~m})$ for a subject the size of a minke whale, and either calm seas or a large stable platform. The arrow or shaft is retrieved by tether or independent flotation. Use of both is recommended. The biopsy plug collected by this method is sufficient for electrophoretic analysis, once the enzyme and buffer systems have been established.

Using a variety of probes for variable DNA regions (see above), the biopsy sample is ample for numerous DNA investigations (as the same DNA extracted from a single skin plug can be probed repeatedly with different probes). Alternatively, the material would be sufficient for $50-100$ restriction enzyme digestions. We would recommend the use of radio-labelled probes in preference to simply staining restriction fragments (see Section 4). To develop a species specific probe, more material may be required (the amount will depend on the tissue: e.g. $4-5 \mathrm{~g}$ of skin), but material from a single animal would be sufficient for this
purpose. Blood can be collected whole and mixed with a preservative that eliminates nucleases (e.g. 5 volumes blood to one volume $\mathrm{pH} 8,0.05 \mathrm{M}$ EDTA solution saturated with NaCl ). This can be stored at room temperature for short periods ( $1-2$ weeks) and at $-20^{\circ} \mathrm{C}$ for much longer (Amos and Hoelzel, in prep). Other tissues (e.g. skin) should be immersed in $20 \%$ DMSO saturated with salt (or saturated salt solution alone) and stored in the same in the same way.

When possible, blood should be collected in EDTA tubes and spun down at $4^{\circ} \mathrm{C}$ to separate the white cells (which can be collected at the interface between the red cells and the plasma). The white cells should then be resuspended in plasma (upper phase) and mixed equal volume with $20 \% \mathrm{DMSO}$ in saline ( $0.9 \% \mathrm{NaCl}$ ). This should then be cooled slowly (about $1^{\circ} \mathrm{C} /$ minute) and stored in liquid nitrogen. This gives very high yields and high molecular weight DNA. Sufficient material for the construction of a genomic library can be acquired by this method from 20 ml of blood. Tissue stored in formaldehyde cannot be used as the DNA will have been structurally altered. Blood and skin give good results. Liver is a concentrated source of DNA, but degrades quickly. For existing stores of samples kept at $-20^{\circ} \mathrm{C}$ for extended periods, probing studies can be conducted using dermal tissues. We have extracted useful DNA from whale skin that had been kept in cold storage for up to 10 years.

Mitochondrial studies are possible from whole cell DNA preparations from small skin samples if a mtDNA probe is available. It is possible to detect mtDNA using probes cloned from a related species (e.g. Dizon, 1987). Fragments smaller than about 250 base pairs will not be detected by the Southern method (see below), which limits resolution (see Wilson et al., 1985). A more thorough analysis requires fresh, or ample frozen material. Liver and heart tissues give the best results. Material collected from strandings or incidental takes should be frozen immediately on dry ice or in liquid nitrogen. Analysis should be conducted as soon as possible: the requisite time will depend on the tissue and species being sampled (see Lansman et al., 1981).

### 5.2 Laboratory procedures

The least expensive technique in terms of materials is enzyme electrophoresis (especially on horizontal starch gels). The apparatus can be easily built from perspex (plexiglass) and two silver wire leads (see Brewer, 1970; Harris and Hopkinson, 1978). Gels can be prepared from a variety of media (starch, polyacrylamide, agar, etc.) and run by various methods (horizontal, vertical, disc, etc.). For screening large numbers of loci or samples in a population study, horizontal starch gels are usually preferred. This is because starch is inexpensive and non-toxic, gels running $10-30$ lanes can be sliced horizontally 4 times and each slice stained for a different enzyme. Resolution for enzymes is often as good as with other media. Poor resolution can sometimes be improved with polyacrylimide gels or iso-electric focusing (see Hames and Rickwood, 1981). Procedures for running and staining gels are outlined in detail in Brewer (1970), Selander et al. (1971), Harris and Hopkinson (1978) and Conkle et al. (1982).

Samples should be prepared by mincing in chilled grinding buffer ( $1: 1$ by volume), mixed well and then allowed to stand on ice for $15-30$ minutes. The mixture is then centrifuged and the supernatent used for
electrophoresis. Best results are obtained with skin samples when the centrifugation step is omitted.

The first step in nuclear DNA analyses is the extraction of high molecular weight genomic DNA. The following protocol was adapted from standard procedures. Samples are ground in liquid nitrogen into a fine powder and then mixed with a 10 mM Tris ( pH 7.4 ) $/ 1 \%$ SDS $/ 0.1 \mathrm{M} \mathrm{NaCl}$ solution with $10 \%$ (by weight) proteinase K and incubated in a $65^{\circ} \mathrm{C}$ water bath for $1-3$ hours. The solution is extracted with equal volumes of phenol, phenol/chloroform and chloroform. The DNA is then precipitated with $1 / 10$ volume NaOAc and 2.5 volumes $100 \%$ ethanol. If high molecular weight DNA is being prepared for cloning procedures and the development of a species-specific probe, then purification by CsCl gradient separation is recommended (see Maniatis et al., 1982).

If a probe is already available, the DNA (taken up in a Tris/EDTA buffer) is digested with a suitable restriction enzyme and run by electrophoresis on an agarose gel. DNA fragments assort on the gel by molecular weight. The DNA is extracted onto a nitrocellulose (or nylon) filter by the Southern blot method (Southern, 1975). The filter is then probed for bands that include DNA fragments with homology for the DNA sequence of interest. The probe is radioactively labelled so that homologous regions can be visualized on X-ray film.

When making a probe, purified DNA is partially digested into random 20 Kb segments (see Maniatis et al., 1982) and cloned into a lambda or plasmid vector. The resultant DNA 'library' can be probed with a radioactively labelled sequence from another species to identify clones that contain the desired sequence. Substantial homology with the human 'minisatellite' probe (Jeffreys et al., 1985b) has been found for various odontocete and mysticete cetaceans (Hoelzel and Amos, 1988). Other variable regions, such as the spacer sequences within the rDNA repeats, require the isolation of species-specific probes that will hybridise with non-conserved spacer regions, (see Section 4.2). Recently, sequences have been found that appear to be useful for probing hypervariable regions in a wide range of species (see Section 4.3 (iv)). For paternity studies, this may eliminate the need for a species specific probe.

If a mtDNA probe is available, whole cell DNA preparations can be investigated using various restriction enzyme digestions. If mitochondrial DNA is to be separated from nuclear DNA, the mitochondria must be extracted separately. This is accomplished by preparing the sample in a glass-tephlon homogenizer, and separating out a mitochondrial pellet by selective centrifugation (see Lansman et al., 1981). The DNA is then phenol extracted, precipitated, and re-suspended in a sucrose/Tris/EDTA buffer. The mitochondrial and nuclear DNA are separated by CsCl gradient centrifugation, and made visible by ethidium bromide staining. The mtDNA will band approximately 0.5 cm below the nuclear DNA (see Giles $e t$ al., 1980; Lansman et al., 1981). Further purification can be achieved by sucrose gradient separation of the mitochondria and repeat CsCl gradients.

Purified mtDNA can then be digested, incorporated into a vector and used to probe subsequent samples. Alternatively mtDNA could be isolated from each sample, digested and restriction fragments visualized by various methods. DNA taken up on a nitrocellulose filter by the Southern blot method (to be analysed by radioactive probing), will have been run on an agarose gel. Agarose,
however, does not give as good resolution of small fragments as polyacrylimide gels. For this reason digested mtDNA samples are often run on polyacrylimide and the bands made visible either by radioactive end-labelling (e.g. Lansman et al., 1981) or silver staining (Tegelstrom, 1986). As described in Section 3.3, one of the assumptions for formulations that describe genetic distance is that all fragments can be detected. However, this difficulty can be overcome if restriction enzymes that digest the genome into relatively few segments are employed. A new technique that employs short 'polymers': sequences with homology for known regions of the mitochondrial genome, and amplification of the intervening sequence, may allow the detection of substantial variation from small samples without the need for a radiolabelled probe, or digesting purified mtDNA with restriction enzymes (Wrischnik et al., 1987).

After a laboratory is set up to conduct recombinant DNA techniques and the probes have been developed, running costs are in the same range as isozyme work (depending on which restriction enzymes are being used). However, establishing a laboratory requires expensive equipment. The most expensive materials are radio-isotopes, CsCl , and the restriction enzymes. The latter vary substantially in price but some are well within an affordable range for routine screening.

### 5.3 Practical interpretation of results

A number of factors can influence the interpretation of allozyme polymorphisms. Many enzymes are composed of more than one polypeptide. These 'multimeric' enzymes show hybrid bands in the heterozygous condition. When staining intensities are symmetrical between all alleles in the heterozygote, the pattern is easily interpreted (see Harris and Hopkinson, 1978). However, sometimes alternative alleles have different enzymatic activities (and therefore staining intensities). In the extreme case, when an allele has no enzymatic activity, the heterozygote will appear identical to the homozygous genotype for the active allele. The so-called 'null' allele can be detected when the genotype for the inactive allele is homozygous. Also, when a null allele is present there will be an apparent deficiency of heterozygotes in the population compared to Hardy-Weinberg expectations.

Sometimes more than one locus codes for the same protein. In many cases the gene products from the different loci can be easily distinguished. However, confusion can arise if the products of multiple loci overlay one another on the gel. It is sometimes possible to separate the loci by using different substrates, changing the pH of the gel buffer, or selectively inhibiting the activity of some loci (see Selander et al., 1971; Harris and Hopkinson, 1978). Some enzymes, such as esterases which are relatively non-specific, are especially prone to this problem.

Another potential problem comes with the interpretation of bands that result from post-translational changes in the structure of the protein. Harris and Hopkinson (1978) discuss the potential causes of these 'secondary' isozyme patterns. The changes are often by oxidation of sulphydryls, deaminations or acetylations and can occur in-vivo or in-vitro during storage, extraction or electrophoresis. Proper storage and handling can help minimize secondary isozyme effects (see Harris and Hopkinson, 1978).

To establish that allelic patterns are consistent with Mendelian inheritance, it is best to conduct breeding
experiments. As this is clearly impractical with cetaceans, consistency with known biochemical properties of the protein and agreement with Hardy-Weinberg expectations should be established as indirect lines of evidence. Having supported the allelic basis of the allozyme polymorphisms, it is important to establish whether alleles at different loci assort independently. Independence can be determined by comparing allelic distributions for all loci. Many of the statistical analyses described above depend on independence between all pairs of loci. In the absence of selection or pleiotropy (one gene having multiple effects), non-random assortment could indicate linkage (proximity of the loci on the same chromosome) or some form of non-random mating, such as polygyny or inbreeding. If known mother-offspring samples are available, the contribution of linkage disequilibrium to non-random assortment can be estimated by statistical analyses (see Smouse and Neel, 1977; Weir and Cockerham, 1978). In this way the role of non-random mating in the population can be assessed.

The interpretation of DNA fingerprints is in some ways less complicated, as the problems associated with the structure and activity of gene products do not apply. Nevertheless, analysis of individuals of known genetic relatedness is useful in initial stages in order to establish the number of allelic and non-allelic fragments in a fingerprint (see Section 4.3 (ii) for details). If DNA fingerprints are to be used to estimate kinship, it would be useful to establish the variance in bandsharing for a given order relationship $a$ priori by comparing known relatives. Interpretation is greatly simplified for the mitochondrial genome where haplotypes appear to be inherited from the maternal parent exclusively.

Most of the complicating factors are a result of distortion or loss of information when the DNA is run electrophoretically on a gel. Distortion can be controlled if gels and samples are carefully prepared. For example, any salt left in the sample preparation will retard the progress of the DNA through the gel, and give a misleading indication of the molecular weight profile. Information is lost through the inability of the gel to resolve very small fragments of DNA (although polyacrylimide is more efficient than agarose). For mtDNA, when the genomic weight is known, it is possible to check for the loss of bands by adding up the weights of bands visible on the gel. Because the mtDNA molecule is circular, the total should equal the genomic weight.

## 6. SUMMARY AND CONCLUSIONS

A detailed analysis of kinship within and between populations will allow the development of population models that incorporate critical parameters such as dispersal rate and effective population size. These parameters can be estimated by measuring existing levels of variation in geographic populations or social groupings for which there is some a priori criteria for sub-dividing the larger sample (see Section 2). Isozyme variation will provide sufficient detail to identify separate populations only when a very large sample of enzymes is screened. This procedure is cumbersome and time consuming in comparison with techniques that analyse genomic components directly (see Sections 3 and 4).

An understanding of stock differentiation will be greatly facilitated by the determination of breeding behaviour. This requires the collection of samples from known
mother/offspring pairs, and some knowledge of the set of potential fathers. The best data-base will come from long-term investigations that track the movements and associations of known individuals. The most appropriate genetic procedure for pedigree analysis and the determination of breeding system is the analysis of hypervariable minisatellite DNA families (see Section 4). It will be necessary to establish some test case genealogies for each species to assess the allelic nature of banding patterns if minisatellites are to be used to describe population structure. Breeding behaviour can also be inferred by statistically comparing genetic variability within and between sub-classes of the population (see Section 2.9).

For a number of pelagic species such as members of the genus Balaenoptera, the a priori classification of populations will have to depend on artificial geographic divisions. However, given these divisions, it is possible to estimate genetic parameters by varying the definition of group boundaries and comparing levels of genetic variation. We recommend RFLP analysis of the mitochondrial genome and the development of species specific probes for the rDNA gene family for this purpose. Mathematical formulations for the estimation of genetic distance, dispersal rate, and other parameters have been adapted from formulations derived for the interpretation of isozyme variation (see Sections 2, 3 and 4).

The strength of the molecular DNA techniques greatly facilitates the collection of samples. A small skin sample stored at room temperature for up to several weeks (or years frozen) in a salt solution is all that is required. There is no need to restrain, capture, or kill subject animals. Samples can be easily collected with a small dart connected to a firing shaft. For sampling a large number of animals on the open sea, the shaft should be fired from a high-test cross-bow or rifle to maximize range.

In this report we have described currently available molecular techniques for the detection of natural variation in proteins and DNA. At the same time we have presented the available methods of analysis of observed variation. From this survey, we conclude that the analysis of defined genomic components such as the mitochondrial genome; the nuclear rDNA multigene family and a variety of hypervariable minisatellite DNA families are the most reliable means for uncovering and understanding the precise and unambiguous genetic differences between stocks and individuals, and their breeding, social and migratory behaviour. DNA fingerprints can also be used as markers for large-scale mark-recapture census programs. The molecular technology for these analyses is now available and exploitable on a large scale, at manageable costs.

## ACKNOWLEDGEMENTS

We would like to thank Bill Amos, John Hrwood, P.G. Smith, G. Winans, O.A. Ryder, G.P. Donovan and John Barrett for their comments and advice during the preparation of this manuscript.

## REFERENCES

Abramowitz, M. and Stegun, J.A. 1964. Handbook of mathematical functions. Appl. Math. Ser. 55: 1,044.
Aguilar, A. and Nadal, J. 1984. Obtencion de biopsias hipodermicas de cetaceos en libertad. Invest. Pesquera 48(1): 23-9.

Allendorf, F.W. and Utter, F.M. 1979. Population genetics. pp. 407-54. In: D.J. Randall and J.R. Brett (Eds) Fish Physiology. Academic Press, London. 650pp.
Allendorf, F.W., Christiansen, F.B., Dobson, T., Eanes, W.F. and Frydenberg, O. 1979. Electrophoretic variation in large mammals. I. The polar bear, Thalarctos maritimus. Hereditas 91: 19-22.

Allison, A.C. 1964. Polymorphism and natural selection in human populations. Cold Spring Harbor Symp. Quant. Biol. 29: 137-49.
Anderson, S., Bankeir, A.T., Barrell, B.G., De Bruijn, M.H.L. Coulson, A.R., Drouin, J., Eperon, I.C., Neirlich, D.P., Roe, B.A., Sanger, F., Schreier, P.H., Smith, A.J.H., Staden, R. and Young, I.G. 1981. Sequence and organization of the human mitochondrial genome. Nature 290: 457-65.
Antonovics, J. 1971. The effects of a heterogeneous environment on the genetics of natural populations. Am. Scient. 59: 593-9.
Aquadro, C.F. and Greenberg, B.D. 1982. Human mitochondrial DNA variation and evolution: Analysis of nucleotide sequences from seven individuals. Genetics 103: 287-312.
Arnason, U. 1974. Comparative chromosome studies in Cetacea. Hereditas 77: 1-36.
Arnason, U. 1982. Southern blot hybridization in cetaceans, using killer whale restriction fragment as a probe. Hereditas 97: 47-9.
Arnason, U. and Widegren, B. 1984. Different rates of divergence in highly repetitive DNA of cetaceans. Hereditas 101: 171-7.
Arnason, U., Lutley, R. and Sandholt, B. 1980. Banding studies on six killer whales: an account of C-band polymorphism and G-band patterns. Cytogenet. Cell Genet. 28: 71-8.
Arnason, U., Purdom, I.F. and Jones, K.W. 1982. Cetacean molecular hybridization using balaenopterid satellite DNA cRNAs as probes. Hereditas 97: 33-6.
Arnason, U., Lima-de-Faria, A., Granstrom, H. and Isaksson, M. 1977. Analysis of cetacean DNA. II. Localization of 18 S and 28 S ribosomal RNA cistrons in a heavy DNA component. Hereditas 87: 67-76.
Arnheim, N. 1983. Concerted evolution of multigene families. pp. 38-61. In: M. Nei and R.K. Koehn (Eds) Evolution of Genes and Proteins. Sinauer, Sunderland, MA. i-xiii + 331pp.
Avise, J.C. and Lansman, R.A. 1983. Polymorphism of mitochondrial DNA in populations of higher animals. pp. 147-64. In: M. Nei and R.K. Koehn (eds) Evolution of Genes and Proteins. Sinauer, Sunderland, MA. i-xiii + 331pp.
Avise, J.C. and Selander, R.K. 1972. Evolutionary genetics of the cave dwelling fishes of the genus Astyanax. Evolution 26: 1-19.
Avise, J.C., Giblin-Davidson, C., Laerm, J., Patton, J.C. and Lansman, R.A. 1979. Mitochondrial DNA clones and matriarchal phylogeny within and among geographic populations of the pocket gopher, Geomys pinetis. Proc. Nat. Acad. Sci. USA 76: 6,694-8.
Avise, J. C., Patton, J. C. and Azvadro. 1980. Evolutionary genetics of birds. I. Relationships among North American thrush allies. The Auk 97: 135-47.
Ayala, F.J., Valentine, J.W., Hedgecock, D. and Barr, L.G. 1975. Deep-sea asteroids: high genetic variability in a stable environment. Evolution 29: 203-12.
Balakrishnan, V. and Sanghvi, L.D. 1968. Distance between populations on the basis of attribute data. Biometrics 24: 859-65.
Baltimore, D. 1981. Gene conversion: some implications for immunoglobulin genes. Cell 24: 592-4.
Barton, N.H. and Charlesworth, B. 1984. Genetic revolutions, founder effects, and speciation. Ann. Rev. Ecol. Syst. 15: 133-64.
Battaglia, B. and Beardmore, J.A. 1977. Marine Organisms, Genetics, Ecology and Evolution. Plenum Press, New York. 550pp.
Bibb, M.J., Van Etten, R.A., Wright, C.T., Walberg, M.W. and Clayton, D.A. 1981. Sequence and gene organization of the mouse mitochondrial DNA. Cell 26: 167-80.
Bigg, M.A. 1982. Assessment of killer whale (Orcinus orca) stocks, Vancouver Island, British Columbia. Rep. int. Whal. Commn 32: 655-66.
Bigg, M.A., MacAskie, I.B. and Ellis, G. 1976. Abundance and movements of killer whales off eastern and southern Vancouver Island with comments on management. Prelim. Rep. Arctic Biol. Sta., Ste. Anne de Bellevue, Quebec, Canada. 20pp.
Bird, A.P. 1986. CpG-rich islands and the function of DNA methylation. Nature 321: 209-213.
Bonnell, M.L. and Selander, R.K. 1974. Elephant seals: genetic variation and near extinction. Science 184: 908-9.
Borisov, V.I. 1981a. A comparative analysis of electrophoretic spectra in the Antarctic whales. 1. Hemoglobins and myoglobins of three species of baleen whales and sperm whales. Zool. Zh. 60: 438-42.

Borisov, V.I. 1981b. A comparative analysis of electrophoretic spectra in the Antarctic whales. 2. Lactate dehydrogenase of three species of baleen whales and sperm whales. Zool. Zh. 60: 595-600.
Borst, P. 1972. Mitochondrial nucleic acids. Ann. Rev. Biochem. 41: 333-76.
Brewer, G.J. 1970. An Introduction to Isozyme Techniques. Academic Press, New York. 186pp.
Brown, G.G. and DesRosiers, L.J. 1983. Rat mitochondrial DNA polymorphism; sequence analysis of a hypervariable site for insertions/deletions. Nucleic Acids Res. 11: 6,699-708.
Brown, S.D.M. and Dover, G.A. 1981. The organisation and evolutionary progress of a dispersed repetitive family of sequences in widely separated rodent genomes. J. Mol. Biol. 150: 441-466.
Brown, W.M. 1976. Interspecific comparisons of animal mitochondrial DNAs. PhD thesis, California Inst. Tech., Pasadena, CA.
Brown, W.M. 1980. Polymorphism in mitochondrial DNA of humans revealed by restriction enzyme analysis. Proc. Nat. Acad. Sci. USA 77: 3,605-9.
Brown, W.M. 1983. Evolution of mitochondrial DNA. pp. 62-88 In: M. Nei and R.K. Koehn (Eds). Evolution of Genes and Proteins. Sinauer, Sunderland, MA. i-xiii + 331pp.
Brown, W.M. 1985. The mitochondrial genome of animar pp. 95-130. In: R.J. MacIntyre (Ed.) Molecular Evolutionary Gı tics. Plenum Press, New York. i-xxi + 610pp.
Brown, W.M. and Simpson, M.V. 1981. Intra- and interspecific variation in the mitochondrial genome of Rattus norvegicus and Rattus rattus: Restriction enzyme analysis of variant mitochondrial DNA molecules and their evolutionary relationships. Genetics 97: 125-43.
Brown, W.M. and Simpson, M.V. 1982. Novel features of animal mtDNA evolution as shown by sequences of two rat cytochrome oxidase subunit II genes. Proc. Nat. Acad. Sci. USA 79: 3,246-50.
Brown, W.M. and Vinograd, J. 1974. Restriction endonuclease cleavage maps of mitochondrial DNAs. Proc. Nat. Acad. Sci., USA 71: 4,617-21.
Brown, W.M., George, M. and Wilson, A.C. 1979. Rapid evolution of mitochondrial DNA. Proc. Nat. Acad. Sci. USA 76: 1,967-71.
Brown, W.M., Shine, J. and Goodman, H.M. 1978. Human mitochondrial DNA: Analysis of 7S DNA from the origin of replication. Proc. Nat. Acad. Sci. USA 75: 735-9.
Brown, W.M., Prager, E.M., Wang, A. and Wilson, A.C. 1982. Mitochondrial DNA sequences of primates: Tempo and mode of evolution. J. Mol. Evol. 18: 225-39.
Bryant, E.H. 1974. On the adaptive significance of enzyme polymorphisms in relation to environmental variability. Am. Nat. 108: 1-19.
Buth, D.G. and Crabtree, C.B. 1982. Genetic variability and population structure of Castomus santaanat in the Santa Clara drainage. Copeia 1982: 439-44.
Cann, R.L. and Wilson, A.C. 1983. Length mutations in human mitochondrial DNA. Genetics 104: 699-711.
Cann, R.L., Brown, W.M. and Wilson, A.C. 1984. Polymorphic sites and the mechanism of evolution in human mitochondrial DNA. Genetics 106: 479-99.
Carson, H.L. 1959. Genetic conditions which promote or retard the formation of species. Cold Syring Harbor Symp. Quant. Biol. 24: 82-105.
Carson, H.L. 1971. Speciation and the founder principle. Stadler Genetics Symposia 3: 51-70.
Cavalli-Sforza, L.L. and Bodmer, W.F. 1971. The Genetics of Human Populations. Freeman, San Francisco.
Cavalli-Sforza, L.L. and Edwards, A.W.F. 1967. Phylogenetic analysis: models and estimation procedures. Amer. J. Hum. Genetics 19: 233-57.
Chang, D. and Clayton, D.A. 1985. Priming of human mitochondrial DNA replication occurs at the light strand promoter. Proc. Nat. Acad. Sci. USA 82: 351-5.
Chapman, R.W., Stephens, J.C., Lansman, R.A. and Avise, J.C. 1982. Models of mitochondrial DNA transmission genetics and evolution in higher eucaryotes. Genet. Res. 40: 41-57.
Cockerham, C.C. 1969. Variance of gene frequencies. Evolution 23 : 72-84.
Cockerham, C.C. 1973. Analysis of gene frequencies. Genetics 74: 679-700.
Coen, E.S. and Dover, G.A. 1983. Unequal exchanges and the co-evolution of X and Y rDNA arrays in D. melanogaster. Cell 33: 849-855.
Coen, E.S., Carpenter, R. and Martin, C. 1986. Transposable elements generate novel spatial patterns of gene expression in Antirrhinum majus. Cell 47:285-296.

Coen, E.S., Strachan, T. and Dover, G.A. 1982a. The dynamics of concerted evolution of rDNA and histone gene families in the melanogaster species sub-group of Drosophila. J. Mol. Biol. 158: 17-35.
Coen, E.S., Thoday, J.M. and Dover, G.A. 1982b. Rate of turnover of structural variants in the rDNA gene family of Drosophila melanogaster. Nature 295: 564-8.
Conkle, M.T., Hodgkiss, P.D., Nunnally, L.B. and Hunter, S.C. 1982. Starch gel electrophoresis of conifer seeds: a laboratory manual. USDA General Technical Report PSW-64, Pacific Southwest Forest and Range Experiment Station, Berkeley, CA. 18pp.
Crow, J.F. and Kimura, M. 1970. An Introduction to Population Genetics Theory. Harper and Row, New York. i-xiv +591 pp.
Daly, J.C. 1981. Effects of social organization and environmental diversity on determining the genetic structure of a population of the wild rabbit, Oryctalagus cuniculus. Evolution 35: 687-706.
Danielsdottir, A.K., Duke, E.J., Joyce, P. and Arnason, A. 1988. Progress report on the study of genetic variation at enzyme loci in fin whales (Balaenoptera physalus) caught off Iceland and Spain. Paper SC/40/Ba10 presented to the IWC Scientific Committee, May 1988.

Dawid, I.B. and Blacker, A.W. 1972. Maternal and cytoplasmic inheritance of mitochondrial DNA in Xenopus. Devel. Biol. 29: 152-61.
Dayhoff, M.O. 1972. Atlas of Protein Sequence and Structure. Vol. 5. National Biomedical Research Foundation, Silver Spring, MD.
Dizon, A.E. 1987. Mitochondrial DNA and the genetic structure of spinner dolphin stocks and schools. US Department of Commerce, NOAA, NMFS, Southwest Fisheries Center, Report of Activities, May-June.
Dolansky, L. and Dolansky, M.P. 1952. Table of $\log _{2}$. Technical Report 227. Research Laboratory of Electronics, Mass. Inst. of Tech., Cambridge.
Dorsey, E.M., Hoelzel, A.R. and Stern, S.J. 1982. Photographic identification and behavior of minke whales in Puget Sound. Unpublished report to the World Wildlife Fund - US.
Dover, G.A. 1982. Molecular drive: a cohesive mode of species evolution. Nature 299: 111-17.
Dover, G.A. 1986a. Molecular drive in multigene families: how biological novelties arise, spread and are assimilated. Trends in Genetics 2: 159-65.
Dover, G.A. 1986b. How to drive an egg. Monitor: Trends in Genetics 2: 300-01.
Dover, G.A. 1987. DNA turnover and the molecular clock. J. Mol. Evol. 26: 47-58
Dover, G.A. and Flavell, R.B. 1982 (eds) Genome Evolution. Academic Press (Lond.). i-xv +382 pp .
Dover, G.A. and Flavell, R.B. 1984. Molecular coevolution: DNA variation and the maintenance of function. Cell 38: 623-24.
Dover, G.A. and Strachan, T. 1987. Molecular drive in the evolution of the immune superfamily of genes: the initiation and spread of novelty. In: G. Kelsoe and D.H. Schulze (Eds). Evolution and Vertebrate Immunity. Univ. Texas Press. 464pp.
Eickbush, T.H. and Burke, W.D. 1986. The silkmoth late chorion locus. II. Gradients of gene conversion in two paired multigene families. J. Mol. Biol. 190: 357-66.
Eshel, I. and Feldman, M.W. 1970. On the evolutionary effect of recombination. Theor. Popn Biol. 1: 88-100.
Ewens, W.J., Spielman, R.S. and Harris, H. 1981. Estimation of genetic variation at the DNA level from restriction endonuclease data. Proc. Natl Acad. Sci. USA 78: 3,623-7.
Falconer, D.S. 1981. Introduction to Quantitative Genetics 2nd Ed. Longman House, New York. i-viii + 340pp.
Farris, J.S. 1970. Methods for computing Wagner trees. Syst. Zool. 19: 83-92.
Fauron, C.M.-R. and Wolstenholme, D.R. 1976. Structural heterogeneity of mitochondrial DNA molecules within the genus Drosophila. Proc. Nat. Acad. Sci., USA 73: 3,623-7.
Fedoroff, N.V. 1979. On spacers. Cell 16: 697-710.
Ferris, S.D., Sage, R.D. and Wilson, A.C. 1982. Evidence from mitochondrial DNA sequences that common laboratory strains of inbred mice are descended from a single female. Nature 295: 163-5.
Ferris, S.D., Sage, R.D. and Wilson, A.C. 1984. DNA variation and evolution. Nature 309: 285-6.
Ferris, S.D., Sage, R.D, Prager, E.M., Ritte, U. and Wilson, A.C. 1983. Mitochondrial DNA evolution in mice. Genetics 105: 681-721.
Flavell, R.B., O'Dell, M., Sharp, P., Nevo, E. and Beiles, A. 1986. Variation in the intergenic spacer of ribosomal DNA of wild wheat, Triticum dicoccoides, in Israel. Mol. Biol. Evol. 3: 547-358.

Flowerdew, M.W. 1983. The circumboreal barnicle Balanus balanoides and its subpopulations. pp. 75-84. In: G.S. Oxford and D. Rollinson (Eds). Protein Polymorphism: Adaptive and Taxonomic Significance. Academic Press, New York. I-xi + 405pp.
Foltz, D.W. 1981. Genetic evidence for long-term monogamy in a small rodent, Peromyscus poliantus. Am. Nat. 117: 665-75
Foltz, D.W. and Hoogland, J.L. 1981. Analysis of the mating system in the black tailed prairie dog (Cynomys ludovicianus) by likelihood of paternity. J. Mamm. 62: 706-12.
Gartside, D.F., Dessauer, H.C. and Joanen, T. 1977. Genic homozygosity in an ancient reptile (Alligator mississippienis). Biochem. Genet. 15: 655-63.
Gerbi, S.A. 1985. Evolution of ribosomal DNA. pp. 419-87. In: R.J. MacIntyre (Ed.). Molecular Evolutionary Genetics. Plenum Press, New York. i-xxi +610 pp.
Giles, R.E., Blanc, H., Cann, H.M. and Wallace, D.C. 1980. Maternal inheritance of human mitochondrial DNA. Proc. Nat. Acad. Sci., USA 77: 6,715-9.
Gillum, A.M. and Clayton, D.A. 1978. Displacement loop replication initiation sequence in animal mitochondrial DNA exists as a family of discrete lengths. Proc. Nat. Acad. Sci., USA 75: 677-81.
Gojobori, T. and Nei, M. 1984. Concerted evolution of the immunoglobulin VH gene family. Mol. Biol. Evol. 1: 195-212.
Greenberg, B.D., Newbold, J.E. and Sugino, A. 1983. Intraspecific nucleotide sequence variability surrounding the origin of replication in human mitochondrial DNA. Gene 21: 33-49.
Gyllensten, W., Ryman, N., Reuterwall, C. and Dratch, P. 1983. Genetic differentiation in four European subspecies of red deer (Cervus elephus, L.). Heredity 51: 561-80.
Gyllensten, U., Wharton, D. and Wilson, A.C. 1985. Maternal inheritance of mitochondrial DNA during backcrossing of two species of mice. J. Hered. 76: 321-4.
Hames, B.D. and Rickwood, D. 1981. Gel Electrophoresis of Proteins. IRL Press, Oxford. 220pp.
Hanken, J. and Sherman, P.W. 1981. Multiple paternity in Belding's ground squirrel litters. Science 212: 351-3.
Harris, H. 1966. Enzyme polymorphism in man. Proc. Royal Soc. (London) Ser. B 164: 298-310.
Harris, H. and Hopkinson, D.A. 1978. Handbook of Enzyme Electrophoresis in Human Genetics. American Elsevier, New York. 850pp.
Hartl, D.L. 1980. Principles of Population Genetics Sinauer, Sunderland, MA. 488pp.
Hauswirth, W.W., Van de Walle, M.J., Laipis, P.J. and Olivo, P.D. 1984. Heterogeneous mitochondrial DNA D-loop sequences in bovine tissue. Cell 37: 1,001-7.
Hawkins, R.E. and Klimstra, W.D. 1970. Dispersal behavior and population structure in the white-tailed deer. J. Wildlife Manag. 34: 407-19.
Hayashi, J.I., Yonekawa, H., Gotoh, J., Watanabe, J. and Tagashira, Y. 1978. Strictly maternal inheritance of rat mitochondrial DNA. Biochem. Biophys. Res. Comm. 83: 1,032-8.
Hayashida, H. and Miyata, T. 1983. Unusual evolutionary conservation and frequent DNA segment exchange in class I genes of the major histocompatability complex. Proc. Nat. Acad. Sci. USA 80: 2,671-5.
Hendrick, P.W. 1971. A new approach to measuring genetic similarity. Evolution 25: 276-80.
Highton, R. and Larson, A. 1979. The genetic relationships of the salamanders of the genus Plethodon. Syst. Zool. 28: 579-99.
Higuchi, R., Bowman, B., Freiberger, M., Ryder, O.A. and Wilson, A.C. 1984. DNA sequence from the quagga, an extinct member of the horse family. Nature 312: 282-4.
Hoelzel, A.R. and Amos, W. 1988. DNA fingerprinting and 'scientific' whaling. Nature 333: 305.
Hood, L., Campbell, J.H. and Elgin, S.C.R. 1975. The organisation, expression and evolution of antibody genes and other multigene families. Ann. Rev. Genet. 9: 305-49.
Hood, L., Kronenberg, M. and Humkapiller, T. 1985. T-cell antigen receptors and the immunoglobulin supergene family. Cell 40 : 225-30.
Horwood, J.W. 1980. An analysis of Wada's data. Paper SC/32/Mi26 presented to the IWC Scientific Committee, June 1980. (Unpublished).
Hutchison, C.A., Newbold, J.E., Potter, S.S. and Edgell, M.H. 1974. Maternal inheritance of mammalian mitochondrial DNA. Nature 251: 536-8.
International Whaling Commission. 1986. Report of the Workshop on the Status of Right Whales. Rep. int. Whal. Commn (special issue 10): 1-33.

International Whaling Commission. 1987. Thirty-Seventh Report of the International Whaling Commission. International Whaling Commission, Cambridge. i-ix +404 pp .
Jarman, A.P., Nicholls, R.D., Weatherall, D.J. and Clegg, J.B. 1986. Molecular characterisation of a hypervariable region downstream of the human alpha-globin cluster. $E M B O J 5: 1,857-63$.
Jeffreys, A.J. 1979. DNA sequence variants in the gamma, G, gamma A, delta and beta globin genes of man. Cell 18: 1-10.
Jeffreys, A.J., Brookfield, J.F.Y. and Semeonoff, R. 1985a. Positive identification of an immigration test-case using human DNA fingerprints. Nature 317: 818-9
Jeffreys, A.J., Wilson, V. and Thein, S.L. 1985b. Hypervariable 'minisatellite' regions in human DNA. Nature 314: 67-73.
Jeffreys, A.J., Wilson, V., Thein, S.L., Weatherall, D.J. and Ponder, B.A.J. 1986. DNA 'fingerprints' and segregation analysis of multiple markers in human pedigrees. Am. J. Human. Genet. 39: 11-24.
Johnson, M.J., Wallace, D.C., Ferris, S.D., Rattazzi, M.C. and Cavalli-Sforza, L.L. 1983. Radiation of human mitochondrial DNA types analysed by restricton endonuclease cleavage patterns. $J$. Mol. Evol. 19: 255-71.
Johnson, W.E. and Selander, R.K. 1971. Protein variation and systematics in kangaroo rats (genus Dipodomys). Systemat. Zool. 20: 377-405.
Kaplan, N. and Langley, C.H. 1979. A new estimate of sequence divergence of mitochondrial DNA using restriction endonuclease mapping. J. Mol. Evol. 13: 295-304.
Karn, M.N. and Penrose, L.S. 1951. Birth weight and gestation time in relation to maternal age, parity, and infant survival. Ann. Eugenics 16: 147-64.
Katona, S., Baxter, B., Brazier, O., Kraus, S., Perkins, J.S. and Whitehead, H. 1979. Identification of humpback whales by fluke photographs. pp. 33-44. In: H.E. Winn and B.L. Olla (Eds). Behavior of Marine Animals. Vol. 3. Cetaceans. Plenum Press, New York. 605pp.
Kedes, L.H. 1979. Histone genes and histone messengers. Ann. Rev. Biochem. 48: 837-70.
Kendall, M.D. and Stuart, A. 1961. The Advanced Theory of Statistics. Vol. 2. Inference and Relationship. Hafner, New York.
Kimura, M. 1968. Genetic variability maintained in a finite population due to mutational production of neutral and nearly neutral iso-alleles. Genet. Res. 11: 247-69.
King, J.L. 1973. The probability of electrophoretic identity of proteins as a function of amino acid divergence. J. Mol. Evol. 2: 317-22.
King, J.L. and Jukes, T.H. 1969. Non-Darwinian evolution: Random fixation of selectively neutral mutations. Science 164: 788-98.
King, M.C. and Wilson, A.C. 1975. Evolution at two levels in humans and chimpanzees. Science 188: 107-16.
Kulu, D.D. 1972. Evolution and cytogenetics pp. 503-27. In: S.H. Ridgeway (Ed.). Mammals of the Sea: Biology and Medicine. Thomas Co., Springfield, Ill. i-xiii + 812pp.
Lamb, B.C. and Helmi, S. 1982. The extent to which gene conversion can change allele frequencies in populations. Genet. Res. 39: 199-217.
Landsteiner, K. 1900. Zur Kentinis der antifermentativen lytischen und agglutinierenden Wirkung des Blutserums und der Lymphe. Zentr. Bakteriol. Parasitenk. 27: 357-62.
Lansman, R.A., Avise, J.C. and Huettel, M.D. 1983. Critical experimental test of the possibility of 'paternal leakage' of mitochondrial DNA. Proc. Nat. Acad. Sci. USA 80: 1,969-71.
Lansman, R.A., Shade, R.O., Shapira, J.F. and Avise, J.C. 1981. The use of restriction endonucleases to measure mitochondrial DNA sequence relatedness in natural populations. III. Techniques and potential applications. J. Mol. Evol. 17: 214-26.
Lavigne, D.M., Bogart, J.P., Downer, R.G.H., Danzmann, R.G., Barchard, W.W. and Earle, M. 1978. Genetic uniformity in northwest Atlantic harp seals, Pagophilus groenlandicus. IGNAF Res. Doc. 78/XI/90 Serial no. 5306. 46pp.
Learn, G.H. and Schaal, B.A. 1987. Population subdivision for ribosomal DNA repeat variants in Clematis fremontii. Evol. 41: 433-8.
Levine, H. 1953. Genetic equilibrium when more than one ecological niche is available. Amer. Nat. 87: 331-3.
Levin, D.A. and Crepet, W.L. 1973. Genetic variation in Lycopodium lucidulum: A phylogenetic relic. Evolution 27: 622-32.
Levins, R. 1968. Evolution in Changing Environments. Princeton Univ. Press, Princeton, N.J. 120pp.
Lewontin, R.C. 1972. The apportionment of human diversity. Evol. Biol. 6: 381-98.

Lewontin, R.C. 1974. The Genetic Basis of Evolutionary Change. Columbia University Press, New York. 380pp.
Lewontin, R.C. and Hubby, J.L. 1966. A molecular aproach to the study of genic heterozygosity in natural populations: II. Amount of variation and degree of heterozygosity in natural populations of Drosophila pseudoobscura. Genetics 54: 595-605.
Li, W.-H. 1983. Evolution of duplicate genes and pseudogenes. pp. 14-37. In: M. Nei and R.K. Koehn (Eds). Evolution of Genes and Proteins. Sinauer, Sunderland, MA. i-xiii +331 pp.
Li, W.-H., Gojobori, T. and Nei, M. 1981. Pseudogenes as a paradigm of neutral evolution. Nature 292: 237-9.
Li, W.-H., Lou, C.-C. and Wu, C.-I. 1985a. Evolution of DNA sequences. pp. 1-94 In: R.M. MacIntyre (Ed.) Molecular Evolutionary Genetics. Plenum Press, New York. v-xxi +610 pp.
Li, W.-H., Wu, C.-I. and Lou, C.-C. 1985b. A new method for estimating anonymous and non-synonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and condon changes. Mol. Biol. Evol. 2: 150-174.
Long, E.H. and Dawid, I.B. 1980. Repeated genes in eukaryotes. Ann. Rev. Biochem. 49: 727-61.
Lynch, M. 1988. Estimates of relatedness by DNA fingerprinting. Mol. Biol. Evol. 5(5): 584-99.
Maniatis, T., Fritsch, E.F. and Sambrook, J. 1982. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, New York. i-x +545 pp.
Maxam, A.M. and Gilbert, W. 1977. A new method for sequencing DNA. Proc. Nat. Acad. Sci. USA 74: 560-4.
Mayr, E. 1963. Animal Species and Evolution. Harvard University Press, Cambridge, MA. 590pp.
McCauley, D.E. and O'Donnell, R. 1984. The effect of multiple mating on genetic relatedness in larval aggregations of the imported willow leaf beetle. Behav. Ecol. Sociobiol. 15: 287-92.
McCracken, G.F. and Bradbury, J.W. 1977. Paternity and heterogeneity in the polygenous bat Phyllostomus hastatus. Science 198: 303-6.
McCracken, G.F. and Bradbury, J.W. 1981. Social organization and kinship in the polygenous bat Phyllostomus hastatus. Behav. Ecol. Sociobiol. 8: 11-34.
Miklos, G.L.G. 1985. Localized highly repetitive DNA sequences. pp. 241-313. In: R.J. MacIntyre (Ed.) Molecular Evolutionary Genetics. Plenum Press, New York. i-xiii +610 pp.
Miyata, T., Yasunaga, T. and Nishida, T. 1980. Nucleotide sequence divergence and functional constraint in mDNA evolution. Proc. Nat. Acad. Sci., USA 77: 7,328-332.
Monnat, R.J. and Loeb, L.A. 1985. Nucleotide sequence preservation of human mitochondrial DNA. Proc. Nat. Acad. Sci., USA 82: 2,895-9.
Moss, T., Mitchelson, K. and de Winter, R. 1985. The promotion of ribosomal transcription in eukaryotes. Oxford Survey Euk. Genes 2: 207-50.
Nagylaki, T. and Petes, T.D. 1982. Intrachromosomal gene conversion and the maintenance of sequence homogeneity among repeated genes. Genetics 100: 315-37.
Nakamura, Y., Leppert, M., O'Connell, P., Wolff, R., Holm, T., Culver, M., Martin, C., Fujimoto, E., Hoff, M., Kumlin, E. and White, R. 1987. Variable number of tandem repeat (VNTR) markers for human gene mapping. Science 235: 1,616-22.
Neel, J.V., Salzano, F.M., Junqueira, P.C., Keiter, F. and Maybury-Lewis, D. 1964. Am. J. Hum. Genet. 16: 52-140.
Nei, M. 1971. Interspecific gene differences and evolutionary time estimated from electrophoretic data on protein identity. Am. Nat. 105: 385-98.
Nei, M. 1972. Genetic distance between populations. Am. Nat. 106: 283-92.
Nei, M. 1975. Molecular Population Genetics and Evolution. North-Holland, Amsterdam. i-xiii +283 pp.
Nei, M. 1977. F-statistics and analysis of gene diversity in subdivided populations. Ann. Hum. Genet. Lond. 41: 225-32.
Nei, M. 1978. Estimation of heterozygosity and genetic distance from a small number of individuals. Genetics 89: 583.
$\mathrm{Nei}, \mathrm{M}$. 1983. Genetic polymorphism and the role of mutation in evolution. pp. 165-90. In: M. Nei and R.K. Koehn (Eds) Evolution of Genes and Proteins. Sinauer, Sunderland, MA. i-xiii + 331pp.
Nei, M. 1987. Molecular Evolutionary Genetics. Columbia University Press, New York. i-x +512 pp.
Nei, M. and Chakraborty, R. 1973. Genetic distance and electrophoretic identity of proteins between taxa. J. Mol. Evol. 2: 323-8.
Nei, M. and Chakraborty, R. 1976. Electrophoretically silent alleles in a finite population. J. Mol. Evol. 8: 381-5.

Nei, M. and Feldman, M. W. 1972. Identity of genes by descent within and between populations under migration and mutation pressures. Theor. Pop. Biol. 3: 460-5.
Nei, M. and Li, W.-H. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Nat. Acad. Sci., USA 76: 5,269-73.
Nei, M. and Roychoudhury, A.K. 1974a. Sampling variances of heterozygosity and genetic distance. Genetics 76: 379-90.
Nei, M. and Roychoudhury, A.K. 1974b. Genetic variation within and between the three major races of man, Caucasoids, Negroids, and Mongoloids. Am. J. Human. Genet. 26: 421.
Nei, M. and Roychoudhury, A.K. 1982. Genetic relationship and evolution of human races. Evol. Biol. 14: 1-60.
Nei, M. and Tajima, F. 1981. DNA polymorphism detectable by restriction endonucleases. Genetics 97: 145-63.
Nei, M., Maruyama, T. and Chakraborty, R. 1975. The bottleneck effect and genetic variability in populations. Evolution 29: 1-10.
Nevo, E. 1978. Genetic variation in natural populations: patterns and theory. Theor. Pop. Biol. 13: 121-77.
Nevo, E. 1979. Adaptive convergence and divergence of subterranean mammals. A. Rev. Ecol. Syst. 10: 269-308.
Nevo, E. 1983. Genetic resouces of wild emmer wheat: structure, evolution and application in breeding. pp. 421-43. In: S. Sakamoto (Ed.) Proceedings of the Sixth International Wheat Genetics Symposium. Kyoto University Press.
Nevo, E. and Shaw, C.R. 1972. Genetic variation in a subterranean mammal, Spalax ehrenbergi. Biochem. Genet. 7: 235-41.
Nevo, E., Beilles, A. and Ben-Shlomo, R. 1983. The evolutionary significance of genetic diversity: Ecological, demographic and life history correlates. Lecture Notes in Biomathematics. Springer, Berlin.
Nevo, E., Kim, Y.J., Shaw, C.R. and Thaeler, C.S. 1974. Genetic variation, selection, and speciation in Thomomys talpoides. Evolution 28: 1-23.
Nevo, E., Golenberg, E., Beiles, A., Brown, A.H.D. and Zohary, D. 1982. Genetic diversity and environmental associations of wild wheat, Triticum dicoccoides. Isr. Theor. Appl. Genet. 62: 241-154.
Ohta, T. 1980. Evolution of multigene families. Lecture Notes on Biomathematics. Springer, Berlin. 90pp.
Ohta, T. 1983. On the evolution of multigene families. Theor. Pop. Biol. 23: 216-40.
Ohta, T. and Dover, G.A. 1984. The cohesive population genetics of molecular drive. Genetics 108: 501-21.
Osborne, R.W., Felleman, F.L. and Heimlich-Boran, J.R. 1985. Some variations in the behavior and ecology of socially isolated pods of killer whales (Orcinus orca). Unpublished sub-contract report to Cascadia Research Collective, Olympia, WA. NOAA Contract No. 82-ABC-00221.
Pamilo, P. 1982. Multiple mating in Formica ants. Hereditas 97: 37-45.
Pamilo, P. 1984. Genotype correlation and regression in social groups: multiple alleles, multiple loci and subdivided populations. Genetics 107: 107-20.
Pamilo, P. and Crozier, R.H. 1982. Measuring genetic relatedness in natural populations: methodology. Theor. Pop. Biol. 21: 171-93.
Patton, J.L. and Feder, J.H. 1981. Microspacial genetic heterogeneity in pocket gophers: non-random breeding and drift. Evolution 31: 697-720.
Patton, J.P. and Yang, S.Y. 1977. Genetic variation in Thomomys bottae pocket gophers: macrogeographic patterns. Evolution 31: 697-720.
Payne, R.S. 1972. Report from Patagonia: The right whales. Unpublished report to the New York Zoological Society.
Perrin, W.F. 1975. Distribution and differentiation of populations of dolphins of the genus Stenella in the eastern tropical Pacific. J. Fish. Res. Bd Can. 32: 1,059-67.
Porter, F.L. and McCracken, G.F. 1983. Social behavior and allozyme variation in a captive colony of Carollia perspicillata. J. Mammal. 64: 295-8.
Powell, J.R. 1975. Protein variation in natural populations of animals. Evol. Biol. 8: 79-119.
Prevosti, A. 1955. Geographical variability in quantitative traits in populations of Drosophila subobscura. Cold Spring Harbor Symp. Quant. Biol. 20: 294-9.
Quinn, T.W., Quinn, J.S., Cooke, F. and White, B.N. 1987. DNA marker analysis detects multiple maternity and paternity in single broods of the lesser snow goose. Nature 326: 393-4.
Reeder, R.H. 1984. Enhancers and ribosomal gene spacers. Cell 38: 349-351.
Rendel, J.M. 1953. Heterosis. Am. Nat. 87: 129-38.

Robberson, D.L., Clayton, D.A. and Morrow, J.F. 1974. Cleavage of replicative forms of mitochondrial DNA by Eco RI endonuclease. Proc. Nat. Acad. Sci. USA 71: 4,447-51.
Rogers, J.S. 1972. Measures of genetic similarity and genetic distance. Studies in Genetics. Vol. VII (Univ. of Texas Publ. No. 7213). pp.145-53.
Sanger, F., Nicklen, S. and Coulson, A.R. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Nat. Acad. Sci. USA 74: 5,463-7.
Sanghvi, L.D. 1953. Comparison of genetical and morphological method for a study of biological differences. Am. J. Phys. Anthrop. 11: 385-404.
Schaller, S.B. 1972. The Serengeti Lion: A study of Predator-Prey Relationships. University of Chicago Press, Chicago. 350pp.
Schwartz, O.A. and Armitage, K.B. 1980 . Genetic variation in social mammals: the marmot model. Science 207: 665-7.
Selander, R.K. 1976. Genetic variation in natural populations. pp. 21-45. In: F. Ayala (Ed.) Molecular Evolution. Sinauer, Sunderland, MA. i-x + 277pp.
Selander, R.K., Hunt, W.G. and Yang, S.Y. 1969. Protein polymorphism and genetic heterozygosity in two European subspecies of the house mouse. Evolution 23: 379-90.
Selander, R.K., Kaufman, D.W., Baker, R.J. and Williams, S.L. 1975. Genic and chromosomal differentiation in pocket gophers of the Geomys busarius group. Evolution 28: 557-64.
Selander, R.K., Yang, S.Y., Lewontin, R.C. and Johnson, W.E. 1970. Genetic variation in the horseshoe crab (Limulus polyphemus), a phylogenetic 'relic'. Evolution 24: 402-14.
Selander, R.K., Smith, M.H., Yang, S.V., Johnson, W.E. and Gentry, J.B. 1971. Biochemical polymorphism and systematics in the genus Peromyscus. I. Variation in the old field mouse. Stud. Genet. Univ. Texas Publ. No. 7103 6: 49-90.
Sharp, G.D. 1975. A study of the electrophoretic and oxegen dissociation properties of the hemoglobins of some delphinid cetaceans. Comp. Biochem. Physiol. (A Com. Physiol.) 51: 673-81.
Sharp, G.D. 1976. Biochemical genetic studies and their value and limitations in stock identification and discrimination of pelagic mammal species. In: FAO Fisheries Series, No. 5, Vol. III, Mammals in the Seas. Rome.
Shah, D.M. and Langley, C.H. 1979. Inter- and intraspecific variation in restriction maps of Drosophila mitochondrial DNAs. Nature 281: 696-9.
Sherman, P.W. and Morton, M.L. 1979. Matrilocal dispersal behavior of the Belding's ground squirrel. Nat. Hist 88: 50-7.
Shimura, E. and Numachi, K.-I. 1987. Genetic variability and differentiation in the toothed whales. Sci. Rep. Whales. Res. Inst. 38: 141-63.
Sibley, C.G. and Ahlquist, J.E. 1984. The phylogeny of the hominoid primates, as indicated by DNA-DNA hybridisation. J. Mol. Evol. 20: 2-15.
Simonsen, V., Born, E.W. and Kistensen, T. 1982a. Electrophoretic variation in large mammals. IV. The Atlantic walrus, Odobenus rosmarus rosmarus (L.). Hereditas 97: 91-4.
Simonsen, V., Kapel, F. and Larsen, F. 1982b. Electrophoretic variation in the minke whale, Balaenoptera acutorostrata Lacepede. Rep. int. Whal. Commn 32: 275-78.
Singer, M.F. 1982. Highly repeated sequences in mammalian genomes. Int. Rev. Cytol. 76: 67-112.
Smith, D.G. 1980. Paternity exclusion in six captive groups of rhesus monkeys (Macaca mulatta). Am. J. Phys. Anthro. 53: 243-9.
Smith, G.P. 1974. Unequal crossover and the evolution of multigene families. Cold Spring Harb. Symp. Quant. Biol. 38: 507-513.
Smith, P.J. and Fujio, Y. 1982. Genetic variation in marine teleosts: high variability in habitat specialists and low variability in habitat generalists. Marine Biol. 69: 7-20.
Smithies, O. and Powers, P.A. 1986. Gene conversions and their relation to homologous chromosome pairing. (In: B.C. Clarke, A. Robertson and A.J. Jeffreys (Eds) Evolution of DNA Sequences). Phil. Trans. Roy. Soc. B. 312: 291-302.
Smouse, P.E. and Chakraborty, R. 1986. The use of restriction fragment length polymorphisms in paternity analysis. Ann. J. Hum. Genet. 38: 918-939.
Smouse, P.E. and Neel, J.V. 1977. Multivariate analysis of gametic disequilibrium in the Yanomama. Genetics 85: 733-52.
Smouse, P.E., Speilman, R.S. and Park, M.H. 1982. Multiple-locus allocation of groups as a function of the genetic variation within and differences among human populations. Am. Nat. 119: 445-63.
Sokal, R.R. and Sneath, P.H.A. 1963. Principles of Numerical Taxonomy. W.H. Freeman and Co., San Francisco.

Sokal, R.R. and Rolf, F.I. 1969. Biometry. The Principles and Practice of Statistics in Biological Research. Freeman and Co., San Francisco. 776pp.
Southern, E. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98: 503.
Southern, S.O., Southern, P.J. and Dizon, A.E. In prep. Molecular and phylogenetic studies with a cloned dolphin mitochondrial genome. J. Mol. Evol.
Strachan, T., Coen, E.S., Webb, D.A. and Dover, G.A. 1982. Modes and rates of change of complex DNA families of Drosophila. J. Mol. Biol. 158: 37-54.
Strachan, T., Webb, D.A. and Dover, G.A. 1985. Transition stages during molecular drive in multiple copy DNA families in Drosophila. EMBO J. 4: 1,701-08.
Swofford, D.L. and Selander, R.K. 1981. BIOSYS-1: a FORTRAN program for the comprehensive analysis of electrophoretic data in population genetics and systematics. J. Hered. 72: 281-3.
Szostack, J.W. and Wu, R. 1980. Unequal crossing over in the ribosomal DNA of Saccharomyces cerevisiae. Nature 284:426-430.
Tanhauser, S.M. 1985. Evolution of mitochondrial DNA: Patterns and rate of change. PhD thesis, University of Florida, Gainesville, FL.
Tautz, D., Trick, M. and Dover, G.A. 1986. Cryptic simplicity in DNA is a major source of genetic variation. Nature 322: 652-6.
Tautz, D., Tautz, C., Webb, D.A. and Dover, G.A. 1987. Evolutionary divergence of promoters and spacers of the rDNA family of four Drosophila species: implications for molecular coevolution in multigene families. J. Mol. Biol. 195: 525-42.
Tegelstrom, H. 1986. Mitochondrial DNA in natural populations: an improved routine for the screening of genetic variation based on sensitive silver staining. Electrophoresis 7: 226-9.
Templeton, A.R. 1977. Survival probabilities of mutant alleles in fine-grained environments. Am. Nat. 111: 951-66.
Templeton, A.R. and Rothman, E.D. 1978. Evolution and fine-grained environmental runs. In: Hooker, Leach and McClennan (Eds). Foundations and Applications of Decision Theory, Vol. II. D. Reidel, Dordrecht, Holland.
Testa, J.W. 1986. Electromorph variation in Weddell seals (Leptonychotes weddelli). J. Mamm. 67(3): 606-10.
Tolliver, D.K., Smith, M.H. and Leftwich, R.H. 1985. Genetic variability in Insectivora. J. Mamm. 66(2): 405-10.
Upholt, W.B. 1977. Estimation of DNA sequence divergence from comparison of restriction endonuclease digests. Nucleic Acids Res. 4: 1,257-65.
Upholt, W.B. and Dawid, I.B. 1977. Mapping of mitochondrial DNA of individual sheep and goats: rapid evolution in the D-loop region. Cell 11: 571-83.
Van Beek, J.G. and de la Mare, W.K. 1981. A statistical review of the biochemical evidence for minke whale stock identity in the Antarctic. Rep. int. Whal. Commn 31: 114-5.
Van Beek, J.G. and Van Biezen, J.B. 1982. A review of the morphological and biochemical research on population identification of southern minke whales. Rep. int. Whal. Commn 32: 753-8.
Van Valen, L. 1965. Morphological variation and the width of the ecological niche. Amer. Nat. 99: 377-90.
Vassart, G., Georges, M., Monsieur, R., Brocas, H., Lequarre, A.S. and Christophe, D. 1987. A sequence of M13 Phage detects hypervariable minisatellites in human and animal DNA. Science 235: 683-4.
Vergnaud, G. 1987. Random synthetic minisatellite probes detect new polymorphic loci in humans. (Pers. comm.).
Wada, S. 1982. Analysis of the biochemical data by G-statistic. Rep. int. Whal. Commn 32: 707.
Wada, S. 1983a. Genetic structure and taxonomic status of minke whales in the coastal waters of Japan. Rep. int. Whal. Commn 33: 361-3.
Wada, S. 1983b. Genetic heterozygosity in the striped dolphin off Japan. Rep. int. Whal. Commn 33: 617-9.
Wada, S. 1984. A note on the gene frequency differences between minke whales from Korean and Japanese coastal waters. Rep. int. Whal. Commn 34: 345-7.

Wada, S. and Numachi, K. 1979. External and biochemical characters as an approach to stock identification for the Antarctic minke whale. Rep. int. Whal. Commn 29: 421-32.
Wahlund, S. 1928. Composition of populations and apparent correlations from the standpoint of animal breeding. Hereditas 11: 65-106.
Walberg, M.W. and Clayton, D.A. 1981. Sequence and properties of the human KB cell and mouse L cell D -loop regions of mitochondrial DNA. Nucleic Acids Res. 9(20): 5,411-21.
Ward, P.S. 1983. Genetic relatedness and colony organization in a species complex of Ponerine ants. Behav. Ecol. Sociobiol. 12: 285-99.
Ward, R.D. and Galleguillos, R.A. 1978. Protein variation in the plaice, dab and flounder and their genetic relationships. pp. 71-93 In: B. Baraglia and J.A. Beard (Eds) Marine Organisms: Genetics, Ecology and Evolution. Plenum Press, New York.
Webster, T.P., Selander, R.K. and Yang, S.Y. 1972. Genetic variability and similarity in the Anelis lizards at Bimimi. Evolution 26: 523-35.
Weir, B.S. and Cockerham, C.C. 1978. Testing hypotheses about linkage disequilibrium with multiple alleles. Genetics 88: 633-42.
Weitkamp, L.R., Arends, T., Gallango, M.L., Neel, J.V., Schultz, J. and Shreffler, D.C. 1972. The genetic structure of a tribal population, the Yanomama indians. III. Seven serum protein systems. Ann. Hum. Genet. 35: 271-9.
Whitehouse, H.L.K. 1983. Genetic Recombination - Understanding the Mechanism. Wiley, New York.
Widegren, B., Arnason, U. and Akusjarvi, G. 1985. Characteristics of a conserved 1579-bp highly repetitive component in the killer whale, Orcinus orca. Mol. Biol. Evol. 2: 411-19.
Wilkinson, G.S. 1985. The social organization of the common vampire bat. II. Mating system, genetic structure and relatedness. Behav. Ecol. Sociobiol. 17: 123-34.
Wilson, A.C., Carlson, S.S. and White, T.J. 1977. Biochemical evolution. Ann. Rev. Biochem. 46: 573-639.
Wilson, A.C., Cann, R.L., Carr, S.M., George, M., Gyllensten, U.B., Helm-Bychowski, K.M., Higuchi, R.G., Palumbi, S.R., Prager, E.M., Sage, R.D. and Stoneking, M. 1985. Mitochondrial DNA and two perspectives on evolutionary genetics. Biol. J. Linn. Soc. 26: 375-400.
Winn, H.E., Bischoff, W.L. and Tarushi, A.G. 1973. Cytological sexing of Cetacea. Mar. Biol. 23: 343-6.
Wong, Z., Wilson, V., Jeffreys, A.J. and Thein, S.L. 1986. Cloning a selected fragment from a human DNA 'fingerprint': isolation of an extremely polymorphic minisatellite. Nucl. Acid Res. 14: 4,605-17.
Workman, P.L. 1969. The analysis of simple genetic polymorphisms. Human. Biol. 41: 97-114.
Workman, P.L. and Niswander, J.D. 1970. Population studies on southwest indian tribes. II. Local genetic differentiation in the Papago. Am. J. Hum. Genet. 22: 24-49.
Wright, S. 1940. Breeding structure of populations in relation to speciation. Am. Nat. 74: 232-48.
Wright, S. 1951. The genetical structure of populations. Ann. Eugenet. 15: 323-54.
Wright, S. 1965. The interpretation of population structure by F-statistics with special regard to systems of mating. Evolution 19: 395-420.
Wright, S. 1978. Evolution and the Genetics of Populations. Vol. 4. Variability Within and Among Natural Populations. University of Chicago Press, Chicago, IL. 580pp.
Wrischnik, L.A., Higuchi, R.G., Stoneking, M., Erlich, H.A., Arnheim, N. and Wilson, A.C. 1987. Length mutations in human mitochondrial DNA: direct sequencing of enzymatically amplified DNA. Nucleic Acids Res. 15(2): 529-41.
Yang, S.Y. and Patton, J.L. 1981. Genetic variability and differentiation in the Galapagos finches. The Auk 98: 230-42.
Zuckerkandl, E. and Pauling, L. 1965. Evolutionary divergence and convergence in proteins. pp. 97-166. In: V. Bryson and H.J. Vogel (Eds) Evolving Genes and Proteins. Academic Press, New York. i-xxi +629 pp .

## Appendix I

## SUMMARY GUIDE

## Introduction

This appendix offers a non-technical guide to the potential applications of molecular biology in stock management, and draws attention to those techniques which are particularly appropriate for answering questions that have been posed by the Scientific Committee. For each question the reader is referred to appropriate sections of the full report, and ultimately to the publications which are referenced in those sections. In addition, we indicate how the same techniques can be used to provide information about other aspects of whale biology which are of particular relevance to management. The practical aspects of sample collection, storage and manipulation, and the mathematical interpretation of results are presented in detail in Sections 5 and 2 respectively of the full report; they will not be discussed further here.

## Genetic diversity

Genetic variation can be measured at any level, from differences in particular nucleotide sequences to the phenotypically expressed characters that distinguish larger taxonomic units. Each level requires its own method of analysis, though individual methods can often be used at several levels. Variation within and between populations can be measured by differences in the structure of DNA, proteins, or chromosomes, and in morphology and behaviour.

The DNA molecule, which consists of a long chain of paired nucleic acid bases (nucleotides), can be altered in a variety of ways. The best understood, but by no means the most prevalent, involves the substitution of one base for another within the molecule (a point mutation). Point mutations are usually the result of errors in copying during replication (copy error) and of the action of mutagens in the environment. Other types of mutation involve the loss or inversion of a stretch of DNA (which may vary in length from two to several thousand nucleotides) at a given position, and the movement of lengths of DNA from one position to another (Section 1).

Nucleotide substitution rates in nuclear DNA vary from about $10^{-12}$ to $10^{-7}$ per year. The observed level of variation in a population is the result of a gradual accumulation of mutations (countered by chance backward mutations), coupled to the actions of natural selection, genetic drift and molecular drive (see below). Genetic drift is most likely to be a factor in finite populations (less than $10^{6}$ individuals), increasing in importance with decreasing population size. It is a stochastic effect which either brings new mutations to fixation (when they are represented in all individuals) or eliminates them by chance.

Variation within a population is also generated by recombination during meiosis. The effect of these processes on variation has been reviewed by Lewontin (1974). It has long been considered that the creation of new combinations of alleles by recombination would facilitate an organism's capacity to adapt to changing environments (e.g. Carson, 1959).

Single-copy genes (genes for which an individual has only one allele from each parent)
In general, the lowest rates of nucleotide substitution occur in sequences which are translated into polypeptides. Even
within this category rates are highly variable. They are known to range from 0.004 to $2.8 \times 10^{-9}$ substitutions per year (Dayhoff, 1972; Wilson et al., 1977; Li et al., 1985a). The average rate for mammals is $0.88 \times 10^{-9}$ substitutions per non-synonymous site per year (Li et al., 1985a). The substitution rate at synonymous sites (nucleotide positions that can vary without altering the structure coded for) may be quite uniform (Miyata et al., 1980; Hayashida and Miyata, 1983) - about $5.5 \times 10^{-9}$ substitutions per year. However, Li et al. (1985a) suggest greater variations (from 1.7 to $11.8 \times 10^{-9}$ substitutions per year). In either case, the average rate is five times higher than at non-synonymous sites (Section 1.4).

Multigene families (genes for which an individual carries many identical copies)
Many genes in the nuclear genome of eukaryotic animals exist in multiple copies (multigene families). These arise by mechanisms that amplify a particular portion of the genome. For example, by 'unequal crossingover' during meiosis chromosomes exchange sections of homologous sequences such that one gains extra sequences and the other suffers a corresponding loss. When repeated many times this can lead to a 'tandem array' of a particular gene. The rates of divergence in multigene families (such as those coding for histones, immunoglobulins and ribosomal RNAs) is complicated by the fact that a mutation occurring in one member gene can spread to other copies of the family by any one of several DNA turnover mechanisms (e.g. unequal crossingover; see section 1.3). The family, or some subsection of it, becomes homogenised and evolves as a unit (Arnheim, 1983; Dover, 1982). Furthermore, multiple genes that are in a tandem array are often separated by spacers of non-coding DNA which experience weaker functional constraints than the genes themselves. Therefore no strict comparisons can be made between rates of divergence in single-copy genes and multigene families. With this proviso Ohta (1980) has compared rates of divergence within the immunoglobulin multigene family between humans and rabbits. She describes a rate of $0.7 \times$ $10^{-9}$ substitutions per nucleotide site per year in the coding regions, and $1.8 \times 10^{-9}$ for the spacer regions. This rate of divergence may be a consequence of the homogenising effect of DNA turnover, which is known to be operating at rates from $10^{-2}$ to $10^{-4}$ per kilobase (a kilobase is a sequence of a thousand bases) per generation (Coen et al., 1982a, b; Jeffreys et al., 1985b): at least two orders of magnitude faster than the mutation rate (Dover, 1982, 1986; Sections 1.3 and 4.2).

## Hypervariable mini-satellites ('DNA Fingerprints')

The highest known evolutionary rates are in the 'satellite' DNA sequences such as the 'hypervariable mini-satellite' regions which give rise to individual 'DNA fingerprints', described by Jeffreys and co-workers (Jeffreys et al., 1985a). These are long series of repeats of a short, highly conserved 'core-sequence' with variable flanking sequences. Nucleotide substitution rates in these regions are about $2 \times 10^{-7}$, more than an order of magnitude higher than the average for coding regions. More important in terms of quantifiable variation, is the fact that the repeat number of the core-sequence within a given region of the
genome is highly variable. The assessed rate of unequal crossingover and other mechanisms which are responsible for variation in the repeat number is approximately 0.5 $1.5 \times 10^{-4}$ per kilobase per gamete for sequences in human mini-satellites (Jeffreys et al., 1985b; Sections 1.3 and 4.3).

## Mitochondrial DNA (mt DNA)

Another useful genetic component for the analysis of genetic variation is the DNA found in mitochondria. The mitochondrial genome is a circular, double-stranded molecule. Mitochondrial sequence variation has been investigated by a number of methods, but the best resolution has been achieved using restriction enzyme analysis (Brown and Vinograd 1974; Robberson et al., 1974). Restriction enzymes recognize particular DNA sequences, usually four to six bases long, and cleave double-stranded DNA at that point. The use of 10 enzymes, each recognizing a nucleotide sequence of four bases, can resolve mitochondrial genomes that differ by less than $0.05 \%$ (Wilson et al., 1985). Vertebrate mtDNA is cut into about 60 segments by a typical four-base enzyme (e.g. Brown, 1980; Ferris et al., 1982).

Substitution rates for the mitochondrial genome are highly variable but on average they are five to 10 times greater than in single-copy nuclear DNA (Brown et al., 1979). There are noncoding sequences in the mitochondrial genome, though proportionally far fewer than in the nuclear genome. Substitution at these sites occurs at about the same frequency as for synonymous sites in nuclear protein genes. The most variable part of the mitochondrial genome is the region where replication begins (e.g. Walberg and Clayton, 1981; Chang and Clayton, 1985). The substitution rate in this region is estimated to be three to five times the rate found in the remainder of the genome (Aquadro and Greenberg, 1982; Cann et al., 1984).

In general, the mitochondrial genome is considerably more variable than nuclear DNA. Most polymorphism in mammalian mtDNA results from base substitution (e.g. Greenberg et al., 1983) so that variations visualized using restriction enzymes can be interpreted as the result of changes in nucleotide sequences (Section 3.1).

## Mechanisms that create population differences

Genetic differences between populations become established when two or more breeding groups are physically separated either in space or time. Existing differences in the genetic composition of these populations are maintained and amplified by a variety of mechanisms which include selection, genetic drift, and various DNA turnover processes collectively known as molecular drive (Dover, 1982, 1986a; see Section 1.3). They can be reduced by an exchange of individuals between the groups (migration) and stablizing selection.

In sexually reproducing populations, mutations can spread by any of three basic mechanisms: natural selection, genetic drift, and molecular drive. Natural selection is a consequence of differences in the reproductive success of genetically-distinct individuals. It may increase or decrease the genetic difference between populations. Genetic drift is a consequence of the continual stochastic gain-and-loss of gametes and individuals in a population (see Section 2.2). It results in the chance differentiation of populations, but its effect is countered by migration. Relatively low levels of migration (one breeding individual per
generation) are sufficient to counteract the effects of genetic drift (see below and Section 2.7). Molecular drive is a spreading process which arises from the fact that various DNA turnover mechanisms cause a continual gain and loss of variant genes in the lifetime of individuals (Section 1.3). Extra copies of genes (in particular in multigene families) can be passed on to the next generation in which the DNA turnover mechanisms can cause further fluctuations in variant gene copy-number per individual.

## Measures of variations

The analysis of molecular variation began at the turn of the century with studies on blood types in humans (Landsteiner, 1900). However, the more extensive characterization of variation in other species did not begin until 1966 with the application of gel electrophoresis (Lewontin and Hubby, 1966; Harris, 1966). In the first ten years after the technique was introduced, genetic variation at loci coding for proteins was described for nearly 250 species (see reviews by Powell, 1975; Selander, 1976; Nevo, 1978). It became apparent that there is extensive genetic variation in natural populations. To date well over 1,000 species have been investigated (see Nevo et al., 1983). The emphasis in marine species has been on invertebrates (e.g. Battaglia and Beardmore, 1977; Flowerdew, 1983), cod and salmonids (see review by Allendorf and Utter, 1979).

Variation at protein loci is usually measured in terms of polymorphism ( P - the proportion of examined loci which are variable) and heterozygosity ( H - a measure of the degree to which loci are represented in the heterozygous condition, Section 1). Nevo (1978) computed average P \& H values based on 242 plant and animal species. He found that vertebrates were less variable than the average for all species considered, and that mammals were at the low end of that group (see Section 1.5).

Enzyme variation has been described for the minke whale, Balaenoptera acutorostrata (Simonsen et al., 1982a; Wada, 1983a) and the striped dolphin Stenella coeruleoalba (Wada, 1983b). Forty striped dolphins were investigated at 15 presumptive loci and found to have $\mathrm{P}=0.130$ ( $95 \%$ criterion for polymorphism) and $\mathrm{H}=0.021+0.008$. Sixty-four minke whale liver samples were examined for variation at 15 presumptive loci and a $\mathrm{P}=0.095$ was found ( $90 \%$ criterion for polymorphism), $\mathrm{H}=0.046+0.01$. These levels of variability are for the most part substantially less than typically found in most other mammalian species. Sharp $(1975 ; 1976)$ reported consistently low levels of blood protein polymorphisms in several races and species of delphinid cetaceans, as did Borisov (1981a; 1981b) for mysticetes (Sections 1.5 and 2.1).

Caution is necessary in the interpretation of average $P$ values however. The number of animals sampled and the number of loci examined vary greatly between studies as do criteria for polymorphism. It is usual to define a locus as polymorphic if the frequency of the most common allele is $99 \%$ or less, but other criteria are used and not always stated in the published report. Further, when a small number of loci are investigated (e.g. 24; see Nei et al., 1975) the estimate of average heterozygosity is subject to a large standard error (Nei and Roychoudhury, 1974a). In addition, estimates of heterozygosity increase as the number of loci examined is reduced (Nei, 1975). In Nevo's (1978) review, 24 or less loci were investigated in $74 \%$ of
the studies. Nei (1975) recommends that estimates of average heterozygosity be conducted on as many loci as possible, ideally a random sample of the genome. The number of individuals on the other hand, can be as low as 20 (Nei, 1978).

A number of techniques have been developed to assess levels of genetic variation. The degree and type of variation depends very much on what part of the genome is being investigated. Examining variation in proteins (and therefore in the genes that code for them) has been by far the most common technique. The strength of this approach is its emphasis on DNA sequences that are expressed phenotypically. This allows an investigation into the role of selection in the evolution of gene loci. The procedure is also inexpensive and relatively easy to conduct (especially horizontal starch gel electrophoresis). However, protein studies show variation in single copy coding sequences, the most conserved class of DNA. In order to demonstrate differences between populations it is preferable to look at sequences which are more variable.

A number of new recombinant DNA techniques (genetic engineering) are available to analyse the more variable regions of the genome. Often these procedures involve the isolation of a DNA sequence which is then radioactively labelled and used to 'probe' the genome for similar sequences (see Sections 3, 4 and 5). This work usually requires greater expense and technical expertise, but a far higher level of resolution is gained.

## Testing geographically separate populations for genetic distance

## Protein variation

A number of studies have identified protein polymorphisms at a few loci and used chi-square statistics to sort populations by differences at these loci. The main trouble with this kind of analysis is that it describes only a very small part of the overall variation. It is possible that differences between populations at one particular locus reflect a difference in selection pressure or some other effect not directly related to reproductive behaviour. Other loci which are ignored in the analysis may not show any statistical difference between samples.

A more reliable measure would take into account variation at a large number of randomly chosen protein loci. For this purpose a variety of measures of genetic distance (defined as the gene diversity between populations expressed as a function of genotype frequency - Nei, 1972) have been proposed (Sanghvi, 1953; Prevosti, 1955; Sokal and Sneath, 1963; Cavalli-Sforza and Edwards, 1967; Balakrishnan and Sanghvi, 1968; Hendrick, 1971; Nei, 1972; Rogers, 1972). Nei (1972), Rogers (1972) and Wright (1978) discuss the relative benefits of the various methods. In general, a method well suited to comparing genetic distance between populations within a species is that proposed by Nei (1971; 1972; 1975; 1978) (see Sections 2.3-2.6).

The accuracy of Nei's formulations is limited primarily by the proportion of variation that can be detected by electrophoresis, and any variation between loci in the rate of nucleotide substitution. Other formulations are limited by these same problems. Nei suggests that his measure of genetic distance has the advantages of measuring the accumulated number of gene substitutions per locus, and having a linear relation to evolutionary time (assuming a
constant rate of nuclear substitution). Nei (1975) defines three measures of genetic distance, the minimum, standard and maximum measures.

Nei (1975) has applied his measure of standard genetic distance (D) to a number of studies on populations within species, sub-species, and higher taxonomic divisions. Genetic distance between races (or populations) was always less than a few percent. Nine populations of the kangaroo rat Dipodomys ordii (Johnson and Selander, 1971) showed the greatest variation between pairs of populations with $D$ ranging from 0.000 to 0.058 . Genetic distance between human races varied from 0.011 to 0.019 (Nei and Roychoudhury, 1974b). This is equivalent to 55,000 to 95,000 years of reproductive isolation (see Section 2.6). Seven studies comparing subspecies showed a range of D values from 0.004 to 0.351 . The majority were about an order of magnitude higher than genetic distances between races. At all levels there is considerable variation in the estimates of D ; however, on average higher taxanomic divisions had higher values.
The estimates of genetic distance given above are based on codon differences at each locus. This means that a large number of loci must be examined to achieve an approximation that is close to the real value. When comparing local populations of the same species, deviations from the true value are expected to be upward when only a few loci are available for study. One reason for this is that monomorphic loci in these populations will usually have the same allele (Nei, 1975). In any case, these measures are useful as estimates of relative distance because they do not depend on assumptions about evolutionary forces.
Attempts to use measures of genetic distance on data from cetaceans have led to inconclusive results (Horwood, 1980; Van Beck and Van Biezen, 1982-see Sections 2.4 and 2.5). Even these results should be interpreted with caution because only three loci were investigated, which makes distance measures very approximate.

## $m t D N A$ variation

Even when a large number of protein loci are investigated, the low variability typically found in single copy genes will limit the resolution that can be achieved (see Section 2.2). This situation can be improved by utilizing the greater levels of variation found in the mitochondrial genome. For restriction site polymorphisms it is more appropriate to compare the average number of nucleotide differences per restriction site for two randomly chosen sequences, rather than the distance measures derived for protein studies. Nei and Li (1979) refer to this as the index of nucleotide diversity.

Restriction analysis of the mitochondrial genome has been used for a number of studies on variation between conspecific populations (e.g. Upholt and Dawid, 1977; Avise, et al. 1979; Brown et al., 1979). A study on geographic populations of pocket gophers employed both mtDNA and standard isozyme analyses (Avise et al., 1979). Though little variation was seen at the enzyme loci, regional 'clones' were detected by digesting mtDNA with six 5 -base and 6 -base restriction enzymes. Genetic distances tended to increase proportionally between more distant geographic populations. Dizon and co-workers are currently investigating mtDNA variation in the four proposed regional populations of spinner dolphins (Stenella longirostris) in the eastern tropical Pacific, but this work is not yet published.
mtDNA is inherited maternally through the egg cytoplasm (Dawid abd Blacker, 1972; Hutchison et al., 1974; Hayashi et al., 1978; Avise et al., 1979; Giles et al., 1980). Present evidence suggests that there is strict maternal inheritance without 'paternal leakage' (see Lansman et al., 1983; Gyllensten et al., 1985), though this is still an open question (see Chapman et al., 1982; Wilson et al., 1985; Section 3.2).

The questions of maternal inheritance and the apparent haploidy of mtDNA are of central importance. There are about $10^{5}$ mitochondria in a mammalian egg, but only about 50 in the midpiece of the sperm. If the sperm contributes no mitochondria to the subsequent generation and the mtDNA in the egg is homogeneous, then mtDNA will be transmitted as a haploid genome within matrilines. This would make mtDNA a powerful genetic marker for population studies because variation would be affected by mutation mechanisms only, independent of recombination. This will only be the case if there is no intra-individual variation in mtDNA (complete homoplasmy). Heteroplasmy could arise either as a result of mutation or by paternal contribution. There is some indication that this may occur rarely in rats (Brown and DesRosiers, 1983), cattle (Hauswirth et al., 1984), humans (Monnat and Loeb, 1985), and some other vertebrate species. This could have important consequences for the interpretation of genealogies (see discussion in Wilson et al., 1985).

It has been suggested that the mean rate of divergence for the mitochondrial genome over a wide range of taxa is $2 \%$ per million-years (Wilson et al., 1985; see e.g. Brown et al., 1982; Ferris et al., 1983; Higuchi et al., 1984; Tanhauser, 1985). This estimate was derived from studies where evidence on species divergence (e.g. from fossils) was already available. Given this rate, it is possible to estimate the time of divergence between species or sub-populations, or the time elapsed since a bottleneck event (see below). Using Nei's measure of genetic distance, Wilson et al., (1985) describe the simple relation: $\mathrm{t}=0.5 \mathrm{~d}$, where d is the mean pairwise divergence between two populations or species (in percent). Within a species, this is described as the time since two randomly picked individuals shared a common mother (see Table 4 in Wilson et al., 1985). An estimate of the long-term effective population size is given by dividing this quantity by the mean number of years per generation (see Section 3.3).

Bottlenecks can greatly influence the level of mtDNA variability. For example, if a population of diploid animals is reduced to a single breeding pair, they will have four copies of the nuclear genome, but only one transmissable copy of the mitochondrial genome. Assuming homoplasmy for mtDNA and no paternal leakage (see above), variation in mtDNA will be eliminated; whereas significant nuclear variability can be retained despite brief bottlenecks (see Barton and Charlesworth, 1984; Wilson et al., 1985). This could have a dramatic affect on the interpretation of genetic distance between populations as estimated from mtDNA analysis. If a rare genotype was fixed by a founder event in one of the populations being compared, the apparent distance would indicate far greater genetic division than was justified. On the other hand such genotypes would act as particularly convenient stock markers (see below and Section 3.3).

A number of species show low levels of mtDNA variation compared to nuclear DNA, suggesting the possibility of a bottleneck period (e.g. Ferris et al., 1982;

1983; 1984). For example, the anomalously low level of variation in human mtDNA has led to the speculation that a transient bottleneck was involved in the formation of Homo sapiens (Brown, 1980; Johnson et al., 1983).

## Variation in the rDNA gene family

Variation in the repeated series of genes coding for ribosomal RNA's (rDNA) is also very useful for the comparison of defined populations. The utility of the rDNA for this purpose derives from two features. 1. The non-transcribed 'spacer' regions between the coding genes are divided into repeat sequences the copy-number of which may differ between populations. This leads to variation in the length of the 'spacer'. Some spacers, such as those in Drosophila and Xenopus, contain a number of different repeat sequences. 2. Unequal crossingover is known to occur at both levels of repetition, generating variation in the copy number of repeats within the spacer and in the number of repeats of the whole unit (genes plus spacer - See Figs 1 and 2) repeated in the multigene family. When this region is visualised on a gel, the position of a band is indicative of the length of the fragment (which is a reflection of the number of spacer subrepeats) and its intensity is indicative of the number of whole units of a given length (see Section 4.2; Coen et al., 1982a, b).
rDNA has been extensively used as a genetic marker in plants. These studies are described in Flavell et al. (1986) and Learn and Schaal (1987). The average number of electrophoretic alleles per population, the proportion of polymorphic loci per population and levels of genetic diversity (using the measures derived by Nei, 1975; see Section 2) are highly correlated with rDNA variables. In particular, there is a strong correlation between the genetic diversity index and the number of independently occurring spacer lengths.

Since rDNA repeats are a heterogeneous collection, even within a single individual, genetic diversity can be apportioned into within-individual, within-sample area, and within-population components using Shannon's information statistic (Lewontin, 1972; see Section 2.3). Using this measure Learn and Schaal (1987) show that an average individual in their study population contains $65.6 \%$ of the total population diversity; and that the average sample area displays more than $95 \%$ of the total diversity. The rDNA length variant patterns parallel the patterns seen for morphological and allozyme markers in other populations (Section 4.2).

## Hypervariable minisatellites of DNA

In general, the hypervariable minisatellites are most applicable for resolving genetic relatedness between individuals in a breeding group. This would require the molecular examination of known mother-offspring pairs. Further, DNA fingerprints could be employed as individual markers for mark-recapture census studies, as each individual will have a 'DNA fingerprint' that is unlikely to be shared with any other whale in the population.

For the identification of races and populations, use can be made of clones of locus-specific arrays of mini-satellites. One particular array has proven useful in this respect because it is an extremely polymorphic locus (heterozygosity $=97 \%$ ) isolated from a single band in a human DNA fingerprint (Wong et al., 1986). The locus shows extreme length variation due to allelic variation in
the number of repeat units in a given array of repeats and slight differences in the length of the repeat unit, (see below: Pedigree analysis and stock determination).

However, too much variability can hinder the identification of races and populations, because the high rates of generation of new variants in each population will tend to obliterate any population-specific variants that might have differentiated the populations at the time of separation. This is particularly true where the periodicity of unequal crossingover can vary greatly from one to many repeat units (see Section 4.3), hence generating a high range of length variants (copy-number variation) in each array. Loci that show lower levels of polymorphism than this are more appropriate for the distinction of higher taxonomic units, (for example, spacer length variation in the rDNA). Some bands in a DNA fingerprint are more conserved than others. It remains to be seen if these less variable bands can be used to distinguish populations.

## Summary

The genetic distances between populations that can be separated by some a priori distinction can be investigated by a variety of methods. The most powerful are those that examine genomic components directly, such as restriction enzyme analysis of the mitochondrial genome, the spacer regions in the rDNA multigene family, and possibly DNA minisatellites. The last component is more useful for monitoring pedigrees and breeding behaviour within population, requiring the initial examination of defined mother-offspring pairs (see below).

## Separating genetic stocks from a mixed assemblage

Enzyme polymorphisms at a given set of loci can provide only a statistical distinction between populations. That is, although the means for the two populations may differ, there is generally considerable overlap between the two distributions of variant types.

For this reason it is not possible to classify individuals chosen at random using allele frequencies. One solution is to look for a unique allele, or genetic 'marker', that is indicative of a given population. Unfortunately enzyme variation in marine mammals is quite low (see above) and such markers are rare. However, the characteristics of variation in multigene families such as the rDNA region described above, are well suited to this kind of comparison (Section 4.2). If possible, the breeding populations that contribute to the mixed assemblage should be studied independently. In this way rDNA length variants that are representative of different populations can be identified and looked for in samples from the mixed group. If it is not possible to identify and sample separate breeding populations a priori, we would recommend looking for markers in both mtDNA and rDNA variation patterns. Further, samples should be collected from mother-calf pairs. Genetic comparison of females and their calves will provide an internal standard of pairs of individuals who must be from the same breeding population. Analysis of two independent potential marker systems will provide corroboration of the classification derived from either component considered alone.

If variant types can be classified into two or more subsets of the total sample, then separate stocks can be identified and the period since they last interbred estimated (see Section 2.6 and 3.3), although the efficiency of this will depend on the rate of variation in the marker system used and the degree of variation found. If highly variable
genomic components such as mtDNA and rDNA cannot distinguish populations, the populations can be classified as effectively panmictic (randomly breeding). It is still possible that separate breeding populations exist, but that there is sufficient dispersal to effectively eliminate genetic differentiation between them.

## Migration rates

Surprisingly little mixing is necessary to overcome the effects of genetic drift and maintain high genetic similarity between populations. Crow and Kimura (1970) describe the conditions for the establishment of an equilibrium between migration and random drift. The mathematical arguments are presented in Section 2.7. In practical terms, if one or more reproductively active individual migrates between populations per generation, then there will be little local differentiation. The effect will be less pronounced in species that tend to disperse over a short range, because neighbouring populations will tend to be genetically similar.

Therefore it is possible to place a theoretical limit on dispersal levels. If separate populations are known to exist, and there is no detectable differentiation between them, then it is probable that at least one reproductive animal migrates between populations each generation. Further differentiation could result from selection pressure or from a variety of DNA turnover mechanisms (see Section 1.3). In this case a higher level of exchange would be necessary to homogenize populations. Considering only the effect of drift, it is also possible to estimate a maximum level of migration that could have occurred. Appropriate formulations are described in Section 2.7.

## Pedigree analysis and stock determination

In many cases, especially with the social odontocetes, it would be possible to model cetacean population dynamics more realistically if their breeding system and the nature of kinship within social groups were known. Genetic markers revealed by allozyme analysis can be used to determine this but far more powerful techniques are available.

By far the most powerful technique for pedigree analysis and measuring kinship within social groups is the examination of hypervariable minisatellite regions by the so-called DNA 'fingerprinting' technique. Minisatellites derive their name from their existance as relatively short tandem arrays of repeats scattered on all but the sex chromosomes. Hypervariability is detectable as variation in copy-number of repeats at the different loci. This is the same as the high variability in copy-number of subrepeats in each rDNA unit which gives rise to spacer length variation (see Section 4.2 and Fig. 2). As with the rDNA, differences between individuals with respect to the repeat copy-number in each array can be detected with the use of a restriction enzyme that cuts outside of the array and not within any of the repeats. The strength of the various minisatellite techniques (Jeffreys et al., 1985b; Jarman et al., 1986; Vassart et al., 1987) is that the repeat sequences are abundantly distributed around the genome, and each region is highly variable in copy-number. Therefore a single radiolabelled 'probe' for the repeated sequence will give a long series of variable bands on a gel. These bands have been shown to be inherited in a Mendelian manner for human pedigree studies (Jeffreys et al., 1985a). This means that the banding patterns can be used for the direct determination of paternity when mother and offspring are
known (see Section 4.3). The probability of an incorrect indentification can be calculated as approximately $7 \times$ $10^{-22}$ in humans (Jeffreys et al., 1986).

Recent attempts to DNA fingerprint pilot whale, killer whale and grey seal have proved successful (Amos, Hoelzel and Dover, unpublished). To date, the use of the human minisatellite probes described above detect 40-50 polymorphic bands unique to each probe. In grey seals the bands have been shown to be independently inherited in a Mendelian manner, permitting the identification of parent-offspring relationships, (Amos et al., in preparation). Small samples of tissue collected from whales can be adequately preserved in salt-based buffers for several weeks at room temperature and mild refrigeration before DNA extraction.

If a species' breeding system in known, it is possible to infer what sub-set of the population is likely to disperse and how closely related aggregating whales are likely to be. Such information also provides a better basis for establishing the effective population size (the average number of whales in a given population that contribute an equal proportion of genes to the next generation),
although this can also be estimated from existing variation levels (see Section 3.3). This is especially important in reduced stocks which may be threatened with low productivity due to inbreeding.

## Summary

A small sample of skin collected from live, unrestrained whales is sufficient for population studies. These samples provide ample genomic DNA for the analytical procedures we propose. As few as 20 individuals per population can be sampled when there is clear a priori criteria for dividing the populations (such as geographic distance). A larger sample would be useful when an attempt is made to sort stocks from temporary aggregations. We recommend the analysis of mtDNA, rDNA spacer, and hypervariable minisatellite DNA ('fingerprinting') variation to address questions on stock identity. Both mtDNA and rDNA variation will give an indication of the degree to which apparently isolated populations have differentiated genetically. Minisatellite variation is of sufficient resolution to illustrate important behavioural characteristics, such as breeding system.

## Appendix 2

## GLOSSARY

## Allele

One of a series of possible alternative forms of a given gene differing in DNA sequence and affecting the structure and/or function of a single product (RNA and/or protein).

## Allozyme

Allelic forms of an enzyme that can be distinguished by electrophoresis.

## Assortative mating

Non-random mating during sexual reproduction involving a tendency for males of a particular kind to breed with females of a particular kind.

## Bottleneck

Fluctuations in allelic frequencies when a large population passes through a contracted stage and then expands again with an altered genetic composition as a consequence of genetic drift.

## Chromatid

The two daughter strands of a chromosome after replication, joined by a single centromere.

## Chromosome

Structure containing DNA and proteins in the cell nucleus.

## Codon

A triplet of nucleotides in a gene that specify a given amino-acid in a protein.
'control' region
Sequences of DNA usually near the beginning of a gene that regulate the transcription of the gene.

5', 3' controls
The beginning and end of a gene are called $5^{\prime}$ and $3^{\prime}$ relative to the direction of transcription. Control regions can be at either end.

## Copy number

The number of copies of a given gene in a set of chromosomes, see multigene family.

## Crossing over

The exchange of genetic material between chromosomes due to chromosome 'breakage' and reunion.

## Cryptic simplicity

Regions of DNA in which the frequency of given short DNA motifs (e.g. ATAG) is higher than expected by chance and in which several motifs are scrambled one with another, see slippage and pure simplicity.

DNA
Deoxyribonucleic acid, the molecular basis of heredity.
cDNA
Complementary DNA made by reverse transcription of messenger RNA.

## $r D N A$

The genes for several classes of ribosomal RNA molecules that go into the construction of ribosomes, usually in long tandem arrays in the chromosomes.
$m t D N A$
Circles of DNA in the mitochondrion.

## DNA reannealing

Double stranded DNA separates into single strands when heated which reanneal back into double strands when temperature is lowered.

## DNA turnover

Continual gain-and-loss of regions of DNA due to a variety of mechanisms of rearrangement, see gene conversion, unequal crossing over, slippage, transposition.

## Exon

Part of genes carrying genetic code for polypeptides; see intron.

## Fingerprinting

Separation of the DNA of an individual into defined fragments the lengths of which are determined by the spacing of given restriction enzyme sites. Numbers and lengths of fragments form a unique 'DNA fingerprint' for each individual when probed for mini-satellite sequences (c.f.).

## Flanking controlling sequences

See $5^{\prime}, 3^{\prime}$ controls.

## Founder effect

See bottleneck.

## Frameshift mutation

Insertion or deletion of a nucleotide base in an exon such that the genetic code is read in a different frame.

## Gene conversion

The ability of one allele of a gene (or one member gene of a gene family) to alter the sequence of another allele (or another member gene) to its own sequence. Usually occurs during meiotic recombination. For example $A a$ can become $A A$ or $a a$. If conversion is biased there is a preference for $A$ or $a$.

## Gene conversion domain

A stretch of DNA involved in gene conversion which can vary from a few to thousands of base-pairs in length.

## Genetic distance

A measure of the number of allelic substitutions per gene that have occurred during the separate evolution of two populations or species.

## Genetic marker

(i) A sequence of DNA, usually recognisable by a restriction enzyme, that is diagnostic for a given chromosome.
(ii) An allele that can distinguish one population from another within a species.

## Genome

The sum total of all the DNA on a haploid set of chromosomes in the nucleus of an individual, including both coding and non-coding sequences.

## Genomic library

A collection of artificially cloned fragments representative of an individual's genome.

## Hardy-Weinberg equilibrium

The law stating that gene frequencies remain constant from generation to generation in an infinitely large, randomly interbreeding population with no selection, migration or mutation.

## Haplotype

A set of alleles of closely-linked genes that tend to be inherited together, uniquely identifying a chromosome.

## Heteroduplex

Formed when one strand of one DNA duplex invades and displaces one strand of another DNA duplex during meiotic recombination, forming a mixed duplex without correct A-T and G-C pairing all along its length.

## Heteroplasmy

Individuals carrying more than one type of mitochondrial or chloroplast DNA.

## Homogenisation

The process, arising as a consequence of DNA turnover, which ensures that most member genes of a multigene family are very similar in sequence, see molecular drive.

## Homologous

Chromosomes that pair during meiosis and contain the same linear arrangement of genes.

## Homoplasmy

Individuals carrying only one type of mitochondrial or chloroplast DNA.

## Hypervariability

Extreme genetic variation between individuals in certain genomic sequences, see fingerprinting.

## Hybridisation stringencies

The fidelity with which single strands of DNA reanneal depends on the stringency of hybridisation determined by temperature and ionic conditions.

## Intergenic spacer

(IGS) is a region of DNA separating classes of ribosomal RNA genes in tandem arrays.

## Inter-locus variance

Differences between genes in the number and frequency of alleles in a population.

## Internal spacer

See intergenic spacer.

## Intra-locus variance

The frequency distribution of alleles of a gene in a population.

## Intron

Region of DNA which separates exons but which does not code for polypeptides.

## Isozyme

Alternative forms of a single compound enzyme which is composed of polypeptides coded by different genes.

Kilobase
A regioning less of DNA 1000 base pairs in length.

## Linkage disequilibrium

Non-random association of linked genes in the gametes of a population: the tendency of certain alleles of one locus to occur with certain alleles of another locus with frequencies greater than expected by chance.

## Locus

Region of DNA, usually a gene.

## Mini-satellite

Tandem array of from 10 to 50 copies of a non-coding length of DNA. Arrays on different chromosomes are usually with different numbers of repeated copies, giving rise to unique individual DNA fingerprints.

## Mobile elements

Lengths of DNA that can move from one position to another in the genome.

## Molecular drive

A process which spreads mutant genes through a multigene family (homogenisation) and through a sexual population (fixation) as a consequence of a variety of DNA turnover mechanisms in eukaryotic nuclear genomes.

## Monomorphic loci

Genes represented by a single fixed allele in a population.

## Motif

A short defined sequence of DNA or polypeptide.

## Multigene family

A collection of identical or near identical genes in the genome. The numbers of gene copies and their distribution amongst chromosomes varies widely between species depending on the gene family in question.

## Neutral allele

A mutation in a gene that has little or no effect on the reproductive success of the individual carrying the allele.

## Non-genic sequence

The bulk of sequences in nuclear genomes which do not code for polypeptides.

## Non-mendelian segregation

Frequencies of genetic variants amongst the gametes of an individual which are not in accordance with predictions based on Mendel's Laws of Inheritance. All mechanisms of DNA turnover lead to patterns of non-mendelian segregation.

## Nonsense codon

A DNA triplet of bases that does not code for an amino acid but serves as a termination signal during protein translation.

## Non-synonymous substitution

A nucleotide substitution, usually in the first or second position of a codon, that causes a replacement of an amino acid in a polypeptide chain.

## Nucleotide

One of the units (A,T,G,C) from which DNA polymers are formed.

## Point mutation

A mutation involving a single nucleotide substitution.

## Probe

A length of DNA or RNA radioactively labelled used to locate similar sequences in a heterogeneous collection of sequences.

## Pure simplicity

Regions of DNA in which a given short motif (e.g. ATAG) occurs in a tandem array without interruption, see slippage and cryptic simplicity.

## Random mating

Any male mating with a female without preference, see assortative mating.

## Recombinant DNA techniques

Techniques capable of locating, cloning and genetically manipulating genes and other DNA sequences.

## Recombination

The creation of new genetic combinations in progeny by independent assortment or crossing over (c.f.).

## Restriction sites

Short motifs of DNA capable of being recognised by a restriction enzyme leading to the cutting of the DNA molecule into separate fragments. Most restriction enzymes have a unique cutting site.

## RFLP

(Restriction fragment length polymorphism). Mutations that eliminate or create new restriction sites lead to DNA fragments of different lengths amongst individuals.

## Satellite DNA

Long tandem arrays of repeated sequences, usually in millions of copies, generally located at centromeres and telomeres off chromosomes. Generally thought to be generated by unequal crossing over.

## Sibship

Genetic relationships between members of a familial pedigree (e.g. cousins, siblings etc.).

## Single-copy genes

Genes for which only two alleles exist (one from each parent) in a diploid cell.

## Slippage

A mechanism of DNA turnover by which gains-and-losses occur of short motifs (usually less than 10 nucleotides) in a DNA helix leading to pure and cryptic DNA simplicity, (see above).

## Subrepeat

A tandem array of repeats within a larger repeating unit, usually also in tandem array, (e.g. subrepeats within the intergenic spacers of rDNA - see above).

Synonymous substitution
Mutations in a codon, usually at the third position, which do not lead to a change in amino-acid at the polypeptide level.

## Tandem array

Head-to-tail arrangement of repetitive genes or non-coding DNA in the genome.

## Tetrad

The four cellular products resulting from meiosis in a single cell.

## Transposition

See mobile elements. Duplicative transposition occurs when a given DNA region replicates and the extra copy moves to another position in the genome. Non-duplicative transposition occurs when the DNA region moves from one position to another: no extra copies are involved.

Unequal crossing over
Crossing over at a time when two chromatids or two chromosomes are not fully aligned leading to the gain of DNA on one chromatid (chromosome) with an equivalent loss in the other. A process of gain-and-loss which can generate multigene families and maintain their homogeneity, (see DNA turnover, molecular drive).

# Analysis of the Southern Hemisphere Minke Whale Mark-Recovery Data 

S.T. Buckland and E.I. Duff<br>Scottish Agricultural Statistics Service, Aberdeen Unit, Macaulay Land Use Research Institute, Craigiebuckler, Aberdeen AB9 2QJ, Scotland.

Accepted January 1989


#### Abstract

Mark-recovery data on the Southern Hemisphere stocks of the minke whale are evaluated. Analyses are carried out under a number of mark-recapture models, and estimates of abundance and of survival rates are obtained. The major potential sources of bias in the estimates are shown to lead to overestimation of abundance and underestimation of survival rate, and approximate corrections are provided so that estimates can be adjusted if information is available on any of the following: short- and long-term mark-shedding; short- and long-term marking mortality; geographic variation in the probability that a whale is taken; discovery and reporting rates of marks; proportion of marked whales takeable; mis-reporting of numbers effectively marked; inadvertant and unrecorded double-marking of whales. Methods used by previous authors to analyse the data and to assess the above effects are reviewed. Recovery data are found to be too few to yield meaningful estimates of abundance for stock Areas I, II and VI, but estimates are given for the other three Areas. All but one of the bias-adjusted mark-recapture estimates of total size of stocks in the combined Area III $+\mathrm{IV}+\mathrm{V}$ lie within the range 271,000 to 350,000 animals. The one analysis yielding a lower estimate does not utilise the full data set. The relative consistency between estimates arises at least in part because the various analyses make similar assumptions and are based on the same data set. Standard errors of the bias-adjusted estimates are underestimated, since the bias adjustments are assumed to be known. Further underestimation of standard errors arises under the models that assume the population is closed.

Other possible uses of the mark-recovery data are considered briefly, but we conclude that the data are too few, and too uncertain, to allow useful detailed analysis of growth rates or of school structure or integrity. Suggestions for future work, in the light of the recommendations of Pollock (1987), and the method of Hoelzel and Dover (1989) for 'fingerprinting' whales, are made.


## INTRODUCTION

The analysis of mark and recovery data from exploited whale stocks has generated a number of questions, which have been addressed with varying degrees of success by many members of the Scientific Committee of the IWC. The main purpose of this study is to examine the mark-recovery data, and the problems that arise from them, for the Southern Hemisphere minke whale stocks.

We assess the value of the data for estimating abundance, survival rates and movements. The effects of the following factors on the analyses are of particular interest: non-closure of the population; recapture effort that varies over space and time; immediate and long-term mark-shedding; short- and long-term marking mortality; non-detection and non-reporting of marks; identification of whales effectively marked; and heterogeneity arising from segregation with respect to age, sex and time.

We will also review guidelines for possible future mark-recapture studies suggested by Pollock (1987), and consider further studies and data collections that may either help to interpret existing data, or enable some of the difficulties exhibited by existing data to be avoided in any future experiments.

We first consider the categories of mark-recapture model relevant to the analyses of the current data, and briefly review their use to date. We then consider analyses of the data under the different models, and address the problems listed above as they arise. An assessment of possible future studies follows, and finally we give a brief overview of the value and limitations of mark-recapture data for obtaining information on the Southern Hemisphere minke whale stocks.

[^9]
## METHODOLOGY

## The Petersen and Chapman's modified estimators

Suppose a sample of $n_{1}$ whales was marked immediately before the start of the whaling season, and during that season, $n_{2}$ whales were taken, of which $m_{2}$ had been marked. Then the Petersen estimate (or Lincoln index) of population size N is given by:

$$
\hat{\mathrm{N}}_{\mathrm{p}}=\mathrm{n}_{1} \cdot \mathrm{n}_{2} / \mathrm{m}_{2}
$$

In the context of commercial whaling, Chapman's modification of this estimator (Chapman, 1951) will not be unbiased, but will have lower bias than the Petersen estimator, at least when the assumptions of the method hold. In addition, it has a valid variance estimator, which has low bias in most circumstances:

$$
\begin{gathered}
\hat{\mathrm{N}}_{\mathrm{c}}=\left[\left(\mathrm{n}_{1}+1\right)\left(\mathrm{n}_{2}+1\right) /\left(\mathrm{m}_{2}+1\right)\right]-1 \\
\hat{\mathrm{~V}}\left(\hat{\mathrm{~N}}_{\mathrm{c}}\right)=\left[\left(\mathrm{n}_{1}+1\right)\left(\mathrm{n}_{2}+1\right)\left(\mathrm{n}_{1}-\mathrm{m}_{2}\right)\left(\mathrm{n}_{2}-\mathrm{m}_{2}\right)\right] / \\
\\
{\left[\left(\mathrm{m}_{2}+1\right)^{2}\left(\mathrm{~m}_{2}+2\right)\right] .}
\end{gathered}
$$

Seber (1982, p.59) lists the assumptions as follows:
(a) the population is closed, so that N is constant;
(b) all animals have the same probability of capture in the first sample;
(c) marking does not affect the catchability of an animal;
(d) the second sample is a simple random sample;
(e) animals do not lose their marks in the time between the two samples;
(f) all marks are reported on recovery in the second sample.
In fact, if assumptions (a) and (d) hold, we do not require (b) and (c). Similarly, if (a), (b) and (c) hold, we do not require (d). However, to minimise the effects of departures from the assumptions, we should attempt to satisfy each
one. In the presence of either birth/immigration or death/emigration, but not both, assumption (a) is not required, although in the latter case, the probability of dying or emigrating should be independent of whether the animal is marked. These estimators and their assumptions are discussed more fully by Seber (1982, pp.59-61) and Pollock (1987). We concentrate here on the applications of the method in the whaling industry, and the specific difficulties that arise. Note that Chapman's modification to the Petersen estimator is more appropriate than that of Bailey (1951; 1952) in this context, since, for the latter to be valid, we must either approximate the hypergeometric distribution that Chapman assumes by a binomial distribution, or assume that sampling in the second sample is with replacement, when clearly it is not.

Several papers adopt the above two-sample method to analyse mark-recovery data within a single season. These include Best and Butterworth (1980), Butterworth and Best (1982), IWC (1982, pp.740-1) and IWC (1985). This has the advantage that it is more reasonable to assume that the population is closed (i.e. no births or deaths, and no migration between Areas) over a short period, but as noted by IWC (1985, p.79), the assumption that marked whales randomly mingle with the unmarked seems dubious in this case. Also, as noted by Butterworth, there may be some trauma involved in marking, which may affect probability of capture over a short period, although Best considers this unlikely (IWC, 1984, p.83). It seems unsafe to assume that no change in probability of capture occurs in the short term, even if observations of behaviour seem to support the assumption. Some individuals may show a greater change in behaviour than others, and subtle differences in evasion or running behaviour may have an impact on probability of capture. Since the minke whales are chased rather than stalked in the Southern Hemisphere, any departure from the assumption may be unimportant.

A common application of the two-sample method is to use independent information on mortality rates to estimate the number of marked whales alive during each successive whaling season, and hence apply the Petersen estimate or Chapman's modification at each season. This allows us to discard assumption (a) above, but instead assumes that we know the true mortality rate without error. Examples of the approach are Ohsumi (1974 and 1977), Miyashita (1982a and 1983), Tillman and Breiwick (1983) and Miyashita and Kasamatsu (1985). Calculations of the form $M_{i}=\left(M_{i-1}-m_{i-1}\right) e^{-\mu}+R_{i} . s$, where $M_{i}$ is the number of whales marked and in the population at the outset of season $i, m_{i}$ is the number of recoveries in season $i, \mu$ is the instantaneous death rate from natural causes, $\mathrm{R}_{\mathrm{i}}$ is the number of whales effectively marked during season $i$, and $1-\mathrm{s}$ is the probability that a whale dies as a direct result of being marked, are valid when $\mu$ is known, and may be used in conjunction with various mark-recapture models.

Brown and Best (1981) extend Chapman's modified estimate as follows. Combine two seasons of data, so that $n_{1}$ becomes the sum of marked whales available (adjusted for natural mortality) over the two seasons, $n_{2}$ becomes the average take of whales across the two seasons and $m_{2}$ becomes the sum of recoveries. Although they state that variance estimates can be calculated as for the normal modified estimate, it is not clear how biased such estimates might be. Further, the method suffers from the potential problems arising from both non-random mixing of whales over the short term, and migration between Areas over the longer term. Sigurjónsson and Gunnlaugsson (1985)
consider a similar extension of the inverse sampling approach of Chapman (1952), but over a number of seasons, and for the Iceland/West Greenland fin whale population. They provide some discussion of biases as a result of incomplete mixing of stocks.

Horwood (1981) considers estimation of population size when there are no recaptures from a two-sample mark-recapture experiment. He assumes that the probability of catching no whales is $\left(1-n_{1} / \mathrm{N}\right)^{\mathrm{n}_{2}}$, with notation as above. Miyashita (1982b and 1983) and IWC (1982 p.702) adopt the same method. This utilises a binomial approximation to the hypergeometric; the exact probability is given by:

$$
\binom{\mathrm{N}-\mathrm{n}_{1}}{\mathrm{n}_{2}} /\binom{\mathrm{N}}{\mathrm{n}_{2}}
$$

However, given the large size of the Southern Hemisphere minke stocks relative to sample sizes, the difference is trivial. We observe $n_{1}$ and $n_{2}$, and may therefore calculate a population size that corresponds to a fixed probability level of catching no marked whales, as proposed by Horwood (1981). If the assumptions of the method hold, this provides a lower confidence limit for population size, in the sense that the interval from this limit to infinity contains all values $\mathrm{N}_{\mathrm{o}}$ that are not rejected by the null hypothesis $\mathrm{H}_{\mathrm{o}}: \mathrm{N}=\mathrm{N}_{\mathrm{o}}$, when the nominal size of the test is set equal to the chosen probability level. As noted by J.G. Cooke (pers.comm.), the actual size of the test is a function of N , and decreases from the nominal size beyond a certain value for N , so that the method is in practice conservative. As before, consideration should be given to the possibility that probability of catching a whale marked a few days or weeks previously may be different than for an unmarked whale, as a result of either non-random search by the whaling vessels, or change in availability or behaviour of the marked whales. Pollock (1987) recommends that the method not be used, and we concur with this opinion.

## Chapman's multiple sample estimator

Chapman (1952) provided a multiple sample estimator of population size for a stable population. Using the same notation as before, we have $\hat{\mathrm{N}}_{\mathrm{m}}=\left(\Sigma \mathrm{n}_{\mathrm{i}} \mathrm{M}_{\mathrm{i}}\right) /\left(1+\Sigma \mathrm{m}_{\mathrm{i}}\right)=\mathrm{d} / \mathrm{L}$ say, where summation is over all seasons since marking first took place. $\mathrm{M}_{\mathrm{i}}$ may be adjusted for natural mortality, removals of marked whales and new marks using the formula of the previous section. Although Chapman provides a variance formula, several authors (e.g. Tillman and Grenfell, 1980) have used a confidence interval in conjunction with this estimator that is intended for a sequential estimator due to Chapman (1954), details of which are given in Seber (1982, p.189). This sequential estimator is very similar to $\hat{\mathrm{N}}_{\mathrm{m}}$, and we will use the same hybridisation of models here, for consistency and because the approximation is likely to be reasonable. Hence, we have

$$
\hat{\mathrm{V}}\left(\hat{\mathrm{~N}}_{\mathrm{m}}\right)=\hat{\mathrm{N}}_{\mathrm{m}}^{2} / \mathrm{L}
$$

and an approximate $95 \%$ confidence interval of

$$
4 \mathrm{~d} /\{\sqrt{ }(4 \mathrm{~L}-1)+1.96\}^{2}<\mathrm{N}<4 \mathrm{~d} /\{\sqrt{ }(4 \mathrm{~L}-1)-1.96\}^{2} .
$$

Assumptions are as for the previous section, but with the additional assumption that $N$ is constant. In the case of the Petersen estimator, and Chapman's modification of it, we could relax that assumption when a value was assumed for the natural mortality rate. For this reason, Pollock (1987)
prefers to use the unweighted average of (modified) Petersen estimates rather than $\hat{\mathbf{N}}_{\mathrm{m}}$, which he notes is in effect a weighted average. We concur with this view, but note that stock sizes are unlikely to have varied greatly during the relatively short period for which there are recovery data on the Southern Hemisphere minke stocks, so that we might expect the two methods to yield very similar estimates.

The estimator $\hat{\mathbf{N}}_{\mathrm{m}}$ has been used for example by Christensen and Rørvik (1978; 1979; 1980; 1981a), Tillman and Grenfell (1980), Tillman (1981), Tillman and Mizroch (1982), IWC (1982, pp.740-1), Miyashita (1983), Tillman and Breiwick (1983), Beddington, Cooke, Christensen, Øritsland, Øien and Rørvik (1984), Miyashita and Kasamatsu (1985) and Cooke (1986).

## Single-release methods

Since, within an Area, marking of Southern Hemisphere minke whales was carried out largely during a single season, single-release models seem particularly relevant. Although the models of the previous sections have been applied widely in conjunction with a parameter representing natural mortality, the estimation of that parameter from the recovery data seems not to have been investigated. The models described by Seber (1982, pp.256-95) allow such estimation. The model of Parker (1963) was found not to be useful without a further assumption. It was unable to distinguish between the following two cases: (i) the population is stable, with a reasonable natural mortality rate; (ii) the population is increasing, with a very low natural mortality rate. In both cases, if the number of whales taken was roughly constant, the number of marked whales taken would decline over time. In the first case, this is a result of mortality; in the second, the proportion of marked whales is decreasing, even if the number stays more or less constant. To force this model (or other single-release models) to converge to a sensible solution when data are sparse, it is therefore necessary to make a further assumption. We choose here to assume that the population is stable. The assumption implies assumption (c) below. Although it is perhaps unrealistic, the lack of data forces it, or a similar assumption, upon us. We could instead assume that the population would be stable in the absence of whaling operations, but it seems more likely that the population would increase from its current level in this circumstance.

Following the notation of Seber (1982), with simplifications under the above assumption, we have:
$\mathrm{N}=$ size of stock, assumed constant
$\mu=$ instantaneous death rate (natural mortality), assumed constant
$\mathrm{M}_{\mathrm{o}}=$ number of whales successfully marked from this stock
$\mathrm{s}=$ number of subsequent whaling seasons
$\mathrm{M}_{\mathrm{i}}=$ number of marked whales surviving to season i , $i=1, \ldots, s$
$n_{i}=$ number of whales taken in season $i, i=1, \ldots, s$
$\mathrm{m}_{\mathrm{i}}=$ number of marked whales taken during season i , $i=1, \ldots, s$
$\mathrm{a}_{\mathrm{i}}=\mathrm{E}\left[\mathrm{m}_{\mathrm{i}} \mid \mathrm{m}_{1}, \mathrm{~m}_{2}, \ldots, \mathrm{~m}_{\mathrm{i}-1}\right]=\mathrm{n}_{\mathrm{i}} \mathrm{M}_{\mathrm{i}} / \mathrm{N}, \mathrm{i}=1, \ldots, \mathrm{~s}$.
Then, $a_{1}=n_{1} M_{o} e^{-\mu} / N$,

$$
\text { and } a_{i}=\left\{n_{i}\left(M_{o}-\sum_{j=1}^{i-1} m i e^{j \mu}\right) e^{-i \mu}\right\} / N, i=2, \ldots, s
$$

Using a Poisson approximation for the distribution of $\boldsymbol{m}_{i}$ with mean $a_{i}$, we have the likelihood,

$$
\mathrm{L}=\prod_{\mathrm{i}=1}^{\mathrm{s}} \mathrm{e}^{-\mathrm{ai}} \cdot \mathrm{a}_{\mathrm{i}}{ }^{\mathrm{mi}} / \mathrm{m}_{\mathrm{i}}!
$$

The Newton-Raphson technique with a Marquardt procedure now enables us to obtain maximum likelihood estimation of the parameters N and $\mu$, together with variances.

The assumptions of this model, as listed by Seber (1982, p.257), but modified under the assumption of constant $N$, are:
(a) immigration and emigration are negligible;
(b) $\mu$ is constant over time, and irrespective of whether an animal is marked or unmarked;
(c) recruitment is equal to the sum of natural and fishing mortality, so that the population size N is constant;
(d) every whale that is alive at the time of season $i$, whether marked or unmarked, has the same probability of being taken during season i ;
(e) the season is of negligible duration;
(f) marked whales do not lose their marks and all marks are reported on recovery.
We agree with the view of Pollock (1987) that assumption (e) is unlikely to be critical. The assumptions are in effect the same as for the previous model. However, in this case, $\mu$ is estimated internally, and the contribution to the variance of estimates from this source is handled correctly. For Chapman's multiple sample estimator, as generally applied in the whaling literature, $\mu$ is estimated externally, using independent information, and this estimate is then assumed to be the true value. As noted by Pollock (1987), we expect to underestimate variance in this circumstance, so that the variance estimate from Parker's model, modified as above, should be a better indicator of precision than that from Chapman's multiple sample model.

Other models could be tried. For example, that of Chapman (1965) might prove useful. Given the sparsity of data relative to most fishery applications, we do not take the method beyond the above model here.

## Recovery models for survival estimation

A series of models exists for estimation of survival rates from recovery data. These were developed by Brownie, Anderson, Burnham and Robson (1985) for analysing bird banding data. Two computer programs are available from these authors to implement their models: ESTIMATE and BROWNIE. BROWNIE is intended for age-dependent models and to test for sex-specific parameters in adult recovery data. For the Southern Hemisphere minke whale stocks, we have too few data to implement such models. We therefore employ the algorithm ESTIMATE here to fit and test the following models, all of which assume survival is age-independent.

Model 1 assumes that survival rates, catching effort and reporting rates (which can be less than $100 \%$ ) vary from year to year, but are independent of the year the whale was marked. Model 2 is as for model 1, except that survival rate is assumed to be constant over time. Model 3 is a further simplification, and assumes that recovery rate is constant over time, which occurs if both catching effort and reporting rates are constant over time. Brownie et al. (1985) also develop model 0 , which is as for model 1 , except that the recovery rate for the first time period after
marking is assumed to be different from following time periods. If same-season recoveries were included in the analyses, this model would therefore seem to be appropriate. We choose here to exclude same-season recoveries, and do not consider model 0 . Although a case could be made for its use, the relatively sparse recovery data preclude its application here.

The following assumptions are common to the above models:
(a) marked whales are representative of the target population;
(b) there is no mark-shedding;
(c) survival rates are the same for marked and unmarked whales;
(d) the fates of marked whales are independent of each other.

## Jolly-Seber and related models

Pollock (1987) recommends that the Jolly-Seber model, and reduced parameter versions of it, should be applied to the minke whale data. The model is described for example by Seber (1982). It assumes that the population is open, but any emigration from the population should be permanent. As with other mark-recapture models for open populations, it cannot distinguish emigration from death, or whether new animals are immigrants or young. Probability of survival and of capture are assumed to be the same for all animals at any given time, but are allowed to vary from one year to another. Jolly (1982) developed reduced parameter versions, in which either probability of survival or probability of capture, or both, can be assumed to be constant over time. Pollock (1987) discusses these models further.

As Jolly (1965), Seber (1982) and Pollock (1987) point out, at any given sample, the animals released with marks may be different from those that are caught, although a more normal circumstance for this class of models is that the releases correspond to, or at least are a subset of, the animals caught. For whale data on a harvested stock, whether an animal is marked or not can, at present, only be determined if the animal is taken. Therefore, of necessity, whales 'released' (i.e. marked) during season i do not comprise part of the sample of whales caught. (Same-season recoveries are not included in the analyses presented here). Seber (1982) also notes that estimation is possible for the case when recoveries continue after the marking programme has ended. However, although subsequent recoveries may be utilised for estimating parameters corresponding to the period when marking took place, estimation of parameters for the last season in which marking occurred and beyond is not possible without a further assumption. This difficulty is particularly relevant to the Southern Hemisphere minke whale data, since in any given Area, there was at most one release of whales on a scale that allows estimation of population size. We must therefore either combine data across Areas or estimate the number of marked whales surviving to any given season using for example the method described earlier, in which independent information on natural mortality rates is utilised. Assumptions common to the models of this section are:
(a) marked whales are representative of the target population;
(b) probability of capture in season i is the same for all whales;
(c) survival rates are the same for marked and unmarked whales;
(d) there is no mark-shedding;
(e) all marks that are taken are recovered and reported;
(f) the fates of marked whales are independent of each other;
(g) the duration of the season is negligible relative to the time scale over which significant changes in the size of the population can occur.

There is a close correspondence between the recovery models of Brownie et al. (1985) with constant or time-dependent survival and recovery probabilities, and those of Jolly (1982) with constant or time-dependent survival and capture probabilities. The correspondence is particularly close for the minke whale data, since all recaptures are lost on capture, so that they are recoveries as defined by Brownie et al. In fact, it seems that the only difference between them is that the survival parameters for the recovery models represent the probability of neither succumbing to natural mortality nor being taken, whereas under Jolly's models, deaths on capture leave the estimated probabilities of survival unaltered. It follows that we should be able to generate to a close approximation the reduced parameter estimates of Jolly from the ESTIMATE program of Brownie et al., and vice versa.

De la Mare (1985) makes good use of the Jolly-Seber model to analyse data on other species (Southern Hemisphere sei whales and two data sets on Southern Hemisphere fin whales). He concludes that his survival estimates are too small to be believable, and suggests that this is a result of mark-shedding; that is, that the 'survival' rate being estimated is the probability that a marked whale both survives and retains its mark from one season to the next. Chapman (1970) also briefly considers a similar approach to analyse Antarctic fin whale data; otherwise, use of Jolly-Seber type models has been largely unexplored for commercial whaling data, although Garrod and Brown (1980) use a similar method for estimating survival rates.

## ANALYSES

## The data

Japanese and Soviet fishing effort for the seasons 1980/81 to $1986 / 87$ is shown in Figs 1 to 9.


Fig. 1. Location of Japanese (o) and Soviet (x) fishing effort, season 1978/79.


Figs. 2-7. Location of Japanese (o) and Soviet (x) fishing effort, 1979/80-1984/85.


Figs. 8 and 9. Location of Japanese (o) and Soviet (x) fishing effort, 1985/86-1986/87.


Fig. 10. Position of marking of whales marked during 1978/79 to 1983/84. Note that 47 whales marked during 1980/81 were in the Ross Sea section of Area VI.

Table 1
Number of Southern Hemisphere minke whales taken by Area, 1971/72 to 1986/87. Numbers are combined Japanese/Soviet/Brazilian catch. Brazilian catches apply to Area II alone; their 1971 catch is included in the 1971/72 figures, etc

|  | Area |  |  |  |  |  |
| :--- | ---: | ---: | ---: | ---: | ---: | ---: |
| Year | I | II | III | IV | V | VI |
| $1971 / 72$ | 0 | 900 | 354 | 2,644 | 0 | 0 |
| $1972 / 73$ | 0 | 702 | 1,187 | 4,557 | 0 | 0 |
| $1973 / 74$ | 1,257 | 826 | 1,698 | 4,568 | 0 | 13 |
| $1974 / 75$ | 1,870 | 1,567 | 1,359 | 2,230 | 734 | 0 |
| $1975 / 76$ | 1,045 | 2,202 | 2,154 | 881 | 631 | 159 |
| $1976 / 77$ | 943 | 1,641 | 2,876 | 1,600 | 1,467 | 149 |
| $1977 / 78$ | 463 | 1,362 | 1,801 | 963 | 884 | 527 |
| $1978 / 79$ | 653 | 1,087 | 2,496 | 1,370 | 414 | 135 |
| $1979 / 80$ | 944 | 1,040 | 2,708 | 1,861 | 1,344 | 0 |
| $1980 / 81$ | 768 | 1,074 | 1,237 | 2,386 | 1,248 | 429 |
| $1981 / 82$ | 930 | 1,048 | 2,172 | 1,625 | 1,177 | 951 |
| $1982 / 83$ | 694 | 854 | 1,112 | 1,969 | 1,896 | 776 |
| $1983 / 84$ | 624 | 625 | 1,416 | 2,095 | 1,445 | 475 |
| $1984 / 85$ | 624 | 600 | 1,416 | 1,488 | 910 | 530 |
| $1985 / 86$ | 670 | 854 | 1,209 | 1,407 | 1,013 | 414 |
| $1986 / 87$ | 607 | 421 | 1,050 | 1,464 | 1,013 | 414 |
| Total | 12,092 | 16,803 | 26,245 | 33,108 | 14,176 | 4,962 |

Table 2
Number of Southern Hemisphere minke whales marked by Area, 1975/6 to 1983/4. Only definite hits (final verdict $=9$ ) are tabulated

|  | Area |  |  |  |  |  |  |
| :--- | ---: | ---: | ---: | ---: | ---: | ---: | :---: |
| Season | I | II | III | IV | V | VI |  |
| $1975 / 76$ | 1 | 3 | 16 | 0 | 3 | 0 |  |
| $1976 / 77$ | 0 | 0 | $24^{1}$ | 2 | 1 | 6 |  |
| $1977 / 78$ | 0 | 0 | 0 | 0 | 0 | 0 |  |
| $1978 / 79$ | 0 | 0 | 0 | $716^{2}$ | $8^{1}$ | 0 |  |
| $1979 / 80$ | 0 | 0 | 703 | 0 | 0 | 0 |  |
| $1980 / 81$ | 0 | 0 | 0 | 0 | $426^{3}$ | 47 |  |
| $1981 / 82$ | 0 | 471 | 0 | 0 | 0 | 0 |  |
| $1982 / 83$ | 24 | 3 | 0 | 0 | 0 | 0 |  |
| $1983 / 84$ | 0 | 0 | 0 | 0 | 0 | 133 |  |

${ }^{1}$ Includes one same-season recovery.
${ }^{2}$ Includes four same-season recoveries.
${ }^{3}$ Includes two same-season recoveries.

The number of minke whales taken by Area and season is given in Table 1, and numbers marked in Table 2. Some marking data were not available at the time of analysis, so that Table 2 is incomplete. For example, Ivashin (1981a) notes that 52 Southern Hemisphere minke whales were marked under the Soviet marking programme during the 1979/80 season, and Brown and Wada (1982) refer to an unpublished paper by Ivashin (1981b), where details are given of a total of 253 minke whales marked across all Areas between 1957/58 and 1979/80. Table 1 of Wada (1984) also shows additional marks, and some discrepancies, relative to Table 2 here.

The marking locations of whales marked during 1978/79 to 1983/84 are shown in Fig. 10. In Table 3, recorded lengths at marking are shown. Recoveries are listed in full in Table 4, and Table 5 shows Area of marking and of recovery, distances moved and elapsed time between marking and recapture. Table 6 cross-tabulates Area of marking by Area of recovery. Figs 11 to 17 show position of marking and position of recovery of each whale recovered at least one season after marking. Those that had been marked in Area I, II or VI are shown in Fig. 11, those marked in Area III in Fig. 12 (westerly movements only) and Fig. 13 (easterly), Area IV marked whales in Fig. 14 (westerly) and Fig. 15 (easterly), and Area V in Fig. 16 (westerly) and Fig. 17 (easterly). In total, 50 whales had moved in a westerly direction and 35 in an easterly

Table 3
Recorded length at marking of Southern Hemisphere minke whales. Only definite hits (final verdict $=9$ ) are tabulated

|  | Length range in metres |  |  |  |  |  |  |  |
| :--- | ---: | ---: | ---: | ---: | ---: | ---: | ---: | ---: |
| Data set | 4.25 | 6.25 | 6.75 | 7.25 | 7.75 | 8.25 | 8.75 | 9.25 |
|  | -6.25 | -6.75 | -7.25 | -7.75 | -8.25 | -8.75 | -9.25 | -10.25 |
| $1975 / 76$ | 0 | 0 | 1 | 2 | 14 | 6 | 0 | 0 |
| 1976/77 | 0 | 0 | 0 | 2 | 3 | 6 | 17 | 5 |
| IDCR cruise, 1978/79 | 12 | 42 | 71 | 136 | 256 | 142 | 60 | 6 |
| IDCR cruise, 1979/80 | 0 | 2 | 3 | 53 | 444 | 163 | 34 | 1 |
| IDCR cruise, 1980/81 | 0 | 1 | 6 | 30 | 161 | 219 | 50 | 1 |
| IDCR cruise, 1981/82 | 0 | 0 | 1 | 4 | 75 | 278 | 82 | 3 |
| IDCR cruise, $1982 / 83$ | 0 | 0 | 1 | 5 | 51 | 173 | 11 | 0 |
| lDCR cruise, $1983 / 84$ | 0 | 0 | 0 | 5 | 27 | 85 | 9 | 0 |
| Brazilian waters, 1981 | 0 | 0 | 0 | 8 | 10 | 8 | 1 | 0 |
| SW Pacific, 1983 | 0 | 0 | 0 | 0 | 2 | 1 | 0 | 0 |
| Total | 12 | 45 | 83 | 245 | 1,043 | 1,081 | 264 | 16 |



Fig. 11. Position of marking and of recovery of minke whales marked in Areas I and II. No whales have yet been recovered from those marked in Area VI.


Fig. 12. Position of marking and of recovery of minke whales marked in Area III that show a westward movement. Same-season recoveries are not shown.
direction, a difference which is not quite significant at the $10 \%$ level (two-tailed test). Of eight same-season recoveries, six travelled west and two east. Three of the six were marked together and caught on the same day; only one of the three marks could be assigned to a carcase.

Table 2 shows that small numbers of minke whales were marked prior to the main season of marking for each Area. Apart from two same-season recoveries, there have been no recoveries of these whales. For some of the methods considered here, it is not possible to utilise these data. For consistency, we choose to discard them in all analyses; the effects are minor.

If migration from an Area is a result of whales 'drifting' over time, with no particular attachment to any region, we would expect the proportion of marked whales recovered
from outside the Area of marking to increase with time from marking. A chi-square test for trend for seasons one to eight on the data of Table 7 fails to detect such a trend $\left(\chi_{1}^{2}=0.12 ; \mathrm{p}>0.5\right)$. An approximate $95 \%$ confidence interval for the annual rate of change in the proportion of same-area recoveries, which should be negative in the presence of the above effect, is $(-4.0 \%, 5.7 \%)$, so that even if there is 'drift', the dilution rate would seem to be close to zero. The test does not rule out the possibility that most whales remain in the Area of marking, but a proportion drift from one Area to another. However, Table 6 shows that, of 27 recoveries outside the Area of marking, 25 (93\%) were from a contiguous Area, suggesting that the hypothesis is unlikely to be true.
[Text continues on p. 130]

Table 4
Marking and recovery information on all marked Southern Hemisphere minke whales that were recovered before the 1987/88 season. Longitudes and latitudes are rounded down to the nearest degree. $\mathrm{FV}=$ final verdict, $\mathrm{R}=$ recovered by

| Whale no. | Mark no. | Marking record |  |  |  | Recovery record |  |  |  | Whale no. | Mark no. | Marking record |  |  |  | Recovery record |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | Date | Position |  | FV | Date | Position |  |  |  |  | Date | Position |  | FV | Date | Position |  |  |
| 1 | 32344 | 17/02/77 | $61^{\circ} \mathrm{E}$ | 67*S | 8 | 18/02/77 | $59^{\circ} \mathrm{E}$ | $66^{\circ} \mathrm{S}$ | Japan | $47^{2}$ | 35368 | 29/01/80 | $26^{\circ} \mathrm{E}$ | $69^{\circ} \mathrm{S}$ | 9 | 27/12/81 | $30^{\circ} \mathrm{E}$ | $65^{\circ} \mathrm{S}$ | USSR |
| $2^{\prime}$ | 26260 | 31/12/78 | $95^{\circ} \mathrm{E}$ | $63^{\circ} \mathrm{S}$ | 8 | 19/01/80 | $88^{\circ} \mathrm{E}$ | $64^{\circ} \mathrm{S}$ | Japan | 48 | 35395 | 30/01/80 | $24^{\circ} \mathrm{E}$ | $69^{\circ} \mathrm{S}$ | 9 | 14/01/85 | $17^{\circ} \mathrm{E}$ | $68^{\circ} \mathrm{S}$ | USSR |
| $2{ }^{1}$ | 27968 | 31/12/78 | $95^{\circ} \mathrm{E}$ | $63^{\circ} \mathrm{S}$ | 9 | 19/01/80 | $88^{\circ} \mathrm{E}$ | $64^{\circ} \mathrm{S}$ | Japan | $49^{9}$ | 35424 | 30/01/80 | $24^{\circ} \mathrm{E}$ | $69^{\circ} \mathrm{S}$ | 9 | 20/12/82 | $52^{\circ} \mathrm{E}$ | $65^{\circ} \mathrm{S}$ | USSR |
| $3{ }^{2}$ | 27966 | 1/01/79 | $94^{\circ} \mathrm{E}$ | $63^{\circ} \mathrm{S}$ | 8 | 21/01/83 | $109^{\circ} \mathrm{E}$ | $65^{\circ} \mathrm{S}$ | USSR | $49^{\circ}$ | 35425 | 30/01/80 | $24^{\circ} \mathrm{E}$ | $69^{\circ} \mathrm{S}$ | 8 | 20/12/82 | $54^{\circ} \mathrm{E}$ | $65^{\circ} \mathrm{S}$ | USSR |
| $3^{2}$ | 27969 | 1/01/79 | $94^{\circ} \mathrm{E}$ | $63^{\circ} \mathrm{S}$ | 8 | not | recovere |  |  | 50 | 35506 | 31/01/80 | $20^{\circ} \mathrm{E}$ | $69^{\circ} \mathrm{S}$ | 5 | 24/12/84 | $44^{\circ} \mathrm{E}$ | $65^{\circ} \mathrm{S}$ | USSR |
| $3^{2}$ | 27970 | 1/01/79 | $94^{\circ} \mathrm{E}$ | $63^{\circ} \mathrm{S}$ | 9 | 21/01/83 | $109^{\circ} \mathrm{E}$ | $65^{\circ} \mathrm{S}$ | USSR | 51 | 38905 | 31/01/80 | $25^{\circ} \mathrm{E}$ | $67^{\circ} \mathrm{S}$ | 9 | 12/01/87 | $19^{\circ} \mathrm{E}$ | $68^{\circ} \mathrm{S}$ | USSR |
| 4 | 34390 | 5/01/79 | $85^{\circ} \mathrm{E}$ | $65^{\circ} \mathrm{S}$ | 9 | 8/01/80 | $85^{\circ} \mathrm{E}$ | $65^{\circ} \mathrm{S}$ | Japan | $52^{2}$ | 35523 | 1/02/80 | $19^{\circ} \mathrm{E}$ | $69^{\circ} \mathrm{S}$ | 8 | 11/07/82 | $34^{\circ} \mathrm{W}$ | $6{ }^{\circ} \mathrm{S}$ | Brazil |
| 5 | 34501 | 5/01/79 | $88^{\circ} \mathrm{E}$ | $63^{\circ} \mathrm{S}$ | 5 | 27/12/83 | $119^{\circ} \mathrm{E}$ | $64^{\circ} \mathrm{S}$ | Japan | $52^{2}$ | 35524 | 1/02/80 | $19^{\circ} \mathrm{E}$ | $69^{\circ} \mathrm{S}$ | 9 | 11/07/82 | $34^{*}$ W | $6{ }^{\circ} \mathrm{S}$ | Brazil |
| 6 | 34266 | 11/01/79 | $75^{\circ} \mathrm{E}$ | $67^{\circ} \mathrm{S}$ | 9 | 14/01/81 | $74^{\circ} \mathrm{E}$ | $67^{\circ} \mathrm{S}$ | USSR | 53 | 35555 | 1/02/80 | $19^{\circ} \mathrm{E}$ | $68^{\circ} \mathrm{S}$ | 9 | 26/12/81 | $31{ }^{\circ} \mathrm{E}$ | $65^{\circ} \mathrm{S}$ | USSR |
| 7 | 34130 | 12/01/79 | $73^{\circ} \mathrm{E}$ | $68^{\circ} \mathrm{S}$ | 9 | 30/01/79 | $74^{\circ} \mathrm{E}$ | $67^{\circ} \mathrm{S}$ | Japan | $54^{1}$ | 38912 | 1/02/80 | $23^{\circ} \mathrm{E}$ | $67^{\circ} \mathrm{S}$ | 8 | 6/12/82 | $54^{\circ} \mathrm{E}$ | $62^{\circ} \mathrm{S}$ | USSR |
| 8 | 34143 | 12/01/79 | $73^{\circ} \mathrm{E}$ | $67^{\circ} \mathrm{S}$ | 9 | 13/12/83 | $90^{\circ} \mathrm{E}$ | $65^{\circ} \mathrm{S}$ | USSR | $54^{1}$ | 38913 | 1/02/80 | $23{ }^{\circ} \mathrm{E}$ | $67{ }^{\circ} \mathrm{S}$ | 9 | 6/12/82 | $54^{\circ} \mathrm{E}$ | $62^{\circ} \mathrm{S}$ | USSR |
| 9 | 34210 | 12/01/79 | $73^{\circ} \mathrm{E}$ | $67^{\circ} \mathrm{S}$ | 8 | 16/01/87 | $31^{\circ} \mathrm{E}$ | $68^{\circ} \mathrm{S}$ | USSR | 55 | 35596 | 2/02/80 | $18^{\circ} \mathrm{E}$ | $68^{\circ} \mathrm{S}$ | 9 | 8/12/83 | $92^{\circ} \mathrm{E}$ | $63^{\circ} \mathrm{S}$ | Japan |
| 10 | 34216 | 12/01/79 | $73^{\circ} \mathrm{E}$ | $68^{\circ} \mathrm{S}$ | 9 | 19/01/81 | $77^{\circ} \mathrm{E}$ | $68^{\circ} \mathrm{S}$ | USSR | 56 | 35722 | 6/02/80 | $11^{\circ} \mathrm{E}$ | $68^{\circ} \mathrm{S}$ | 9 | 9/12/80 | $56^{\circ} \mathrm{E}$ | $62^{\circ} \mathrm{S}$ | USSR |
| 11 | 34525 | 12/01/79 | $73^{\circ} \mathrm{E}$ | $65^{\circ} \mathrm{S}$ | 9 | 28/01/87 | $76^{\circ} \mathrm{E}$ | $67^{\circ} \mathrm{S}$ | USSR | 57 | 35725 | 6/02/80 | $10^{\circ} \mathrm{E}$ | $68^{\circ} \mathrm{S}$ | 9 | 19/01/82 | $19^{\circ} \mathrm{E}$ | $69^{\circ} \mathrm{S}$ | USSR |
| 12 | 34531 | 12/01/79 | $73^{\circ} \mathrm{E}$ | $65^{\circ} \mathrm{S}$ | 9 | 26/11/82 | $45^{\circ} \mathrm{E}$ | $63^{\circ} \mathrm{S}$ | USSR | 58 | 35837 | 10/02/80 | $5^{\circ} \mathrm{E}$ | $68^{\circ} \mathrm{S}$ | 9 | 6/03/85 | $72^{\circ} \mathrm{E}$ | $66^{\circ} \mathrm{S}$ | USSR |
| 13 | 34559 | 13/01/79 | $71^{\circ} \mathrm{E}$ | $64^{\circ} \mathrm{S}$ | 9 | 27/02/80 | $3^{\circ} \mathrm{E}$ | $69^{\circ} \mathrm{S}$ | Japan | 59 | 35968 | 12/02/80 | $3^{\circ} \mathrm{E}$ | $69^{\circ} \mathrm{S}$ | 9 | 16/01/85 | $14^{\circ} \mathrm{E}$ | $68^{\circ} \mathrm{S}$ | USSR |
| 14 | 34102 | 14/01/79 | $70^{\circ} \mathrm{E}$ | $67^{\circ} \mathrm{S}$ | 9 | 19/12/80 | $59^{\circ} \mathrm{E}$ | $63^{\circ} \mathrm{S}$ | Japan | 60 | 35995 | 13/02/80 | $0^{\circ} \mathrm{E}$ | $69^{\circ} \mathrm{S}$ | 9 | 2/12/81 | $54^{\circ} \mathrm{E}$ | $64^{\circ} \mathrm{S}$ | USSR |
| $15^{3}$ | 34576 | 14/01/79 | $70^{\circ} \mathrm{E}$ | $64^{\circ} \mathrm{S}$ | 9 | 5/03/80 | $71^{\circ} \mathrm{E}$ | $66^{\circ} \mathrm{S}$ | USSR | 61 | 38979 | 26/12/80 | $156^{\circ} \mathrm{E}$ | $63^{\circ} \mathrm{S}$ | 5 | 19/01/86 | $149{ }^{\circ} \mathrm{E}$ | $67^{\circ} \mathrm{S}$ | Japan |
| $? 3$ | 34574 | 14/01/79 | $70^{\circ} \mathrm{E}$ | $64^{\circ} \mathrm{S}$ | 3 | not | recovere | d |  | 62 | 38997 | 27/12/80 | $159{ }^{\circ} \mathrm{E}$ | $63^{\circ} \mathrm{S}$ | 9 | 16/01/85 | $147^{\circ} \mathrm{E}$ | $66^{\circ} \mathrm{S}$ | Japan |
| $?^{3}$ | 34578 | 14/01/79 | $70^{\circ} \mathrm{E}$ | $64^{\circ} \mathrm{S}$ | 3 | not | recovere | d |  | 63 | 43524 | 29/12/80 | $150^{\circ} \mathrm{E}$ | $65^{\circ} \mathrm{S}$ | 9 | 25/01/83 | $142^{\circ} \mathrm{E}$ | $66^{\circ} \mathrm{S}$ | Japan |
| 16 | 33975 | 22/01/79 | $104^{\circ} \mathrm{E}$ | $62^{\circ} \mathrm{S}$ | 9 | 4/01/81 | $81^{\circ} \mathrm{E}$ | $64^{\circ} \mathrm{S}$ | USSR | 64 | 44079 | 30/12/80 | $150^{\circ} \mathrm{E}$ | $64^{\circ} \mathrm{S}$ | 9 | 10/01/84 | $137{ }^{\circ} \mathrm{E}$ | $64^{\circ} \mathrm{S}$ | Japan |
| 17 | 34766 | 24/01/79 | $106{ }^{\circ} \mathrm{E}$ | $64^{\circ} \mathrm{S}$ | 9 | 16/12/81 | $80^{\circ} \mathrm{E}$ | $63^{\circ} \mathrm{S}$ | Japan | 65 | 43551 | 31/12/80 | $147^{\circ} \mathrm{E}$ | $65^{\circ} \mathrm{S}$ | 9 | 23/01/83 | $139^{\circ} \mathrm{E}$ | $66^{\circ} \mathrm{S}$ | Japan |
| 18 | 36822 | 27/01/79 | $110^{\circ} \mathrm{E}$ | $64^{\circ} \mathrm{S}$ | 9 | 9/02/86 | $149^{\circ} \mathrm{E}$ | $65^{\circ} \mathrm{S}$ | Japan | 66 | 39025 | 2/01/81 | $145{ }^{\circ} \mathrm{E}$ | $66^{\circ} \mathrm{S}$ | 9 | 17/01/84 | $148{ }^{\circ} \mathrm{E}$ | $66^{\circ} \mathrm{S}$ | Japan |
| $19^{4}$ | 31461 | 29/01/79 | $116^{\circ} \mathrm{E}$ | $63^{\circ} \mathrm{S}$ | 9 | 5/01/81 | $94^{\circ} \mathrm{E}$ | $63^{\circ} \mathrm{S}$ | Japan | 67 | 39069 | 3/01/81 | $141{ }^{\circ} \mathrm{E}$ | $66^{\circ} \mathrm{S}$ | 9 | 24/01/81 | $142^{\circ} \mathrm{E}$ | $66^{\circ} \mathrm{S}$ | Japan |
| $20^{4}$ | 31466 | 29/01/79 | $116^{\circ} \mathrm{E}$ | $63^{\circ} \mathrm{S}$ | 8 | 5/01/81 | $95^{\circ} \mathrm{E}$ | $63^{\circ} \mathrm{S}$ | Japan | 68 | 39070 | 3/01/81 | $141{ }^{\circ} \mathrm{E}$ | $66^{\circ} \mathrm{S}$ | 9 | 20/01/82 | $108^{\circ} \mathrm{E}$ | $65^{\circ} \mathrm{S}$ | Japan |
| 21 | 36897 | 29/01/79 | $116^{\circ} \mathrm{E}$ | $64^{\circ} \mathrm{S}$ | 5 | 19/12/82 | $124^{\circ} \mathrm{E}$ | $64^{\circ} \mathrm{S}$ | Japan | 69 | 39079 | 3/01/81 | $141{ }^{\circ} \mathrm{E}$ | $66^{\circ} \mathrm{S}$ | 5 | 8/12/82 | $111^{\circ} \mathrm{E}$ | $61^{\circ} \mathrm{S}$ | Japan |
| $22^{5}$ | 36908 | 30/01/79 | $116^{\circ} \mathrm{E}$ | $64^{\circ} \mathrm{S}$ | 8 | 30/12/86 | $137{ }^{\circ} \mathrm{E}$ | $64^{\circ} \mathrm{S}$ | Japan | 70 | 43598 | 3/01/81 | $140^{\circ} \mathrm{E}$ | $66^{\circ} \mathrm{S}$ | 9 | 5/01/87 | $133^{\circ} \mathrm{E}$ | $65^{\circ} \mathrm{S}$ | Japan |
| $22^{5}$ | 36909 | 30/01/79 | $116^{\circ} \mathrm{E}$ | $64^{\circ} \mathrm{S}$ | 9 | 30/12/86 | $137{ }^{\circ} \mathrm{E}$ | $64^{\circ} \mathrm{S}$ | Japan | 71 | 43686 | 3/01/81 | $141{ }^{\circ} \mathrm{E}$ | $66^{\circ} \mathrm{S}$ | 4 | 5/01/82 | $26^{\circ} \mathrm{E}$ | $68^{\circ} \mathrm{S}$ | USSR |
| 235 | 36912 | 30/01/79 | $117^{\circ} \mathrm{E}$ | $64^{\circ} \mathrm{S}$ | 9 | 15/12/82 | $120^{\circ} \mathrm{E}$ | $64^{\circ} \mathrm{S}$ | Japan | 72 | 44105 | 3/01/81 | $142^{\circ} \mathrm{E}$ | $64^{\circ} \mathrm{S}$ | 9 | 17/01/87 | $131{ }^{\circ} \mathrm{E}$ | $65^{\circ} \mathrm{S}$ | Japan |
| $23{ }^{5}$ | 36913 | 30/01/79 | $117^{\circ} \mathrm{E}$ | $64^{\circ} \mathrm{S}$ | 9 | 15/12/82 | $120^{\circ} \mathrm{E}$ | $64^{\circ} \mathrm{S}$ | Japan | 73 | 39127 | 4/01/81 | $140^{\circ} \mathrm{E}$ | $66^{\circ} \mathrm{S}$ | 8 | 20/01/82 | $107{ }^{\circ} \mathrm{E}$ | $65^{\circ} \mathrm{S}$ | Japan |
| 24 | 36931 | 31/01/79 | $120^{\circ} \mathrm{E}$ | $64^{\circ} \mathrm{S}$ | 5 | 18/01/83 | $109^{\circ} \mathrm{E}$ | $65^{\circ} \mathrm{S}$ | USSR | 74 | 39142 | 4/01/81 | $140^{\circ} \mathrm{E}$ | $66^{\circ} \mathrm{S}$ | 8 | 20/01/83 | $139{ }^{\circ} \mathrm{E}$ | $66^{\circ} \mathrm{S}$ | Japan |
| 25 | 36975 | 31/01/79 | $121^{\circ} \mathrm{E}$ | $64^{\circ} \mathrm{S}$ | 9 | 18/12/86 | $111^{\circ} \mathrm{E}$ | $65^{\circ} \mathrm{S}$ | Japan | 75 | 39149 | 4/01/81 | $140^{\circ} \mathrm{E}$ | $66^{\circ} \mathrm{S}$ | 5 | 2/01/83 | $130^{\circ} \mathrm{E}$ | $64^{\circ} \mathrm{S}$ | Japan |
| 26 | 36983 | 31/01/79 | $121^{\circ} \mathrm{E}$ | $64^{\circ} \mathrm{S}$ | 8 | 26/11/82 | $46^{\circ} \mathrm{E}$ | $64^{\circ} \mathrm{S}$ | USSR | 76 | 39227 | 4/01/81 | $140^{\circ} \mathrm{E}$ | $66^{\circ} \mathrm{S}$ | 9 | 13/01/84 | $139{ }^{\circ} \mathrm{E}$ | $65^{\circ} \mathrm{S}$ | Japan |
| 27 | 31447 | 1/02/79 | $121^{\circ} \mathrm{E}$ | $61^{\circ} \mathrm{S}$ | 9 | 30/11/82 | $100^{\circ} \mathrm{E}$ | $61^{\circ} \mathrm{S}$ | Japan | 77 | 39239 | 4/01/81 | $141{ }^{\circ} \mathrm{E}$ | $66^{\circ} \mathrm{S}$ | 9 | 4/01/83 | $131{ }^{\circ} \mathrm{E}$ | $65^{\circ} \mathrm{S}$ | Japan |
| $28^{\circ}$ | 37007 | 1/02/79 | $123^{\circ} \mathrm{E}$ | $64^{\circ} \mathrm{S}$ | 9 | 19/02/79 | $111^{\circ} \mathrm{E}$ | $64^{\circ} \mathrm{S}$ | Japan | 78 | 39253 | 4/01/81 | $141{ }^{\circ} \mathrm{E}$ | $66^{\circ} \mathrm{S}$ | 9 | 18/01/83 | $139{ }^{\circ} \mathrm{E}$ | $66^{\circ} \mathrm{S}$ | Japan |
| $29^{\circ}$ | 37008 | 1/02/79 | $123^{\circ} \mathrm{E}$ | $64^{\circ} \mathrm{S}$ | 9 | 19/02/79 | $111^{\circ} \mathrm{E}$ | $64^{\circ} \mathrm{S}$ | Japan | $79^{5}$ | 39257 | 4/01/81 | $141{ }^{\circ} \mathrm{E}$ | $66^{\circ} \mathrm{S}$ | 7 | 15/01/83 | $138^{\circ} \mathrm{E}$ | $66^{\circ} \mathrm{S}$ | Japan |
| $30^{\circ}$ | 37010 | 1/02/79 | $123^{\circ} \mathrm{E}$ | $64^{\circ} \mathrm{S}$ | 9 | 19/02/79 | $111^{\circ} \mathrm{E}$ | $64^{\circ} \mathrm{S}$ | Japan | $79^{5}$ | 39259 | 4/01/81 | $141{ }^{\circ} \mathrm{E}$ | $66^{\circ} \mathrm{S}$ | 9 | 15/01/83 | $138^{\circ} \mathrm{E}$ | $66^{\circ} \mathrm{S}$ | Japan |
| 31 | 31110 | 2/02/79 | $125^{\circ} \mathrm{E}$ | $61^{\circ} \mathrm{S}$ | 8 | 26/02/86 | $93^{\circ} \mathrm{E}$ | $65^{\circ} \mathrm{S}$ | USSR | 80 | 39272 | 4/01/81 | $141{ }^{\circ} \mathrm{E}$ | $66^{\circ} \mathrm{S}$ | 9 | 27/12/82 | $130^{\circ} \mathrm{E}$ | $64^{\circ} \mathrm{S}$ | Japan |
| 32 | 31117 | 2/02/79 | $125^{\circ} \mathrm{E}$ | $61^{\circ} \mathrm{S}$ | 9 | 5/01/82 | $113^{\circ} \mathrm{E}$ | $64^{\circ} \mathrm{S}$ | Japan | 81 | 39402 | 5/01/81 | $138{ }^{\circ} \mathrm{E}$ | $66^{\circ} \mathrm{S}$ | 9 | 11/02/87 | $144^{\circ} \mathrm{E}$ | $66^{\circ} \mathrm{S}$ | Japan |
| 33 | 37066 | 2/02/79 | $123^{\circ} \mathrm{E}$ | $64^{\circ} \mathrm{S}$ | 9 | 18/03/86 | $114^{\circ} \mathrm{E}$ | $64^{\circ} \mathrm{S}$ | Japan | 82 | 44115 | 6/01/81 | $135^{\circ} \mathrm{E}$ | $63^{\circ} \mathrm{S}$ | 9 | 13/12/86 | $105^{\circ} \mathrm{E}$ | $62^{\circ} \mathrm{S}$ | Japan |
| 34 | 37129 | 4/02/79 | $127^{\circ} \mathrm{E}$ | $64^{\circ} \mathrm{S}$ | 9 | 3/02/87 | $89^{\circ} \mathrm{E}$ | $65^{\circ} \mathrm{S}$ | USSR | 83 | 39176 | 16/01/81 | $161{ }^{\circ} \mathrm{E}$ | $63^{\circ} \mathrm{S}$ | 5 | 3/02/81 | $155^{\circ} \mathrm{E}$ | $65^{\circ} \mathrm{S}$ | Japan |
| 35 | 37142 | 4/02/79 | $127{ }^{\circ} \mathrm{E}$ | $64^{\circ} \mathrm{S}$ | 8 | 19/01/83 | $109{ }^{\circ} \mathrm{E}$ | $65^{\circ} \mathrm{S}$ | USSR | 84 | 44640 | 17/01/81 | $166^{\circ} \mathrm{E}$ | $64^{\circ} \mathrm{S}$ | 9 | 18/02/84 | $166^{\circ} \mathrm{W}$ | $71^{\circ} \mathrm{S}$ | Japan |
| 36 | 37169 | 5/02/79 | $128^{\circ} \mathrm{E}$ | $65^{\circ} \mathrm{S}$ | 9 | 27/12/79 | $89^{\circ} \mathrm{E}$ | $65^{\circ} \mathrm{S}$ | Japan | 85 | 39193 | 21/01/81 | $172{ }^{\circ} \mathrm{E}$ | $68^{\circ} \mathrm{S}$ | 5 | 16/01/87 | $32^{\circ} \mathrm{E}$ | $68^{\circ} \mathrm{S}$ | USSR |
| 37 | 33896 | 6/02/79 | $130^{\circ} \mathrm{E}$ | $65^{\circ} \mathrm{S}$ | 9 | 25/02/79 | $127^{\circ} \mathrm{E}$ | $65^{\circ} \mathrm{S}$ | Japan | $86^{5}$ | 39334 | 25/01/81 | $171{ }^{\circ} \mathrm{E}$ | $72^{\circ} \mathrm{S}$ | 9 | 26/02/83 | $168^{\circ} \mathrm{W}$ | $73^{\circ} \mathrm{S}$ | Japan |
| 38 | 38579 | 31/12/79 | $40^{\circ} \mathrm{E}$ | $67^{\circ} \mathrm{S}$ $62^{\circ} \mathrm{S}$ | 9 | 19/01/86 | $11^{\circ} \mathrm{E}$ | $69^{\circ} \mathrm{S}$ | USSR | $86^{5}$ | 39342 | 25/01/81 | $171{ }^{\circ} \mathrm{E}$ | $72^{\circ} \mathrm{S}$ | 9 | 26/02/83 | $168{ }^{\circ} \mathrm{W}$ | $73^{\circ} \mathrm{S}$ | Japan |
| 39 | 35078 | 2/01/80 | $44^{\circ} \mathrm{E}$ | $62^{\circ} \mathrm{S}$ | 8 | 14/12/84 | $17^{\circ} \mathrm{E}$ | $59^{\circ} \mathrm{S}$ | USSR | 87 | 44781 | 31/01/81 | $177^{\circ} \mathrm{W}$ | $72{ }^{\circ} \mathrm{S}$ | 9 | 3/01/86 | $160^{\circ} \mathrm{W}$ | $64^{\circ} \mathrm{S}$ | Japan |
| 40 | 35170 | 9/01/80 | $60^{\circ} \mathrm{E}$ | $64^{\circ} \mathrm{S}$ | 9 | 16/12/80 | $62^{\circ} \mathrm{E}$ | $63^{\circ} \mathrm{S}$ | Japan | 88 | 43010 | 7/01/82 | $35^{\circ} \mathrm{W}$ | 62*S | 9 | $1 / 09 / 85$ | $34^{\circ} \mathrm{W}$ | $7{ }^{\circ} \mathrm{S}$ | Brazil |
| 417 | 35240 | 12/01/80 | $64^{\circ} \mathrm{E}$ | $6^{63}{ }^{\circ} \mathrm{S}$ | 8 | 8/03/85 | $71{ }^{\circ} \mathrm{E}$ | $66^{\circ} \mathrm{S}$ | USSR | 89 | 43214 | 29/01/82 | $12^{\circ} \mathrm{W}$ | $69^{\circ} \mathrm{S}$ | 9 | ??/01/85 | $5^{\circ} \mathrm{E}$ | $66^{\circ} \mathrm{S}$ | USSR |
| $42^{7}$ 43 | 38820 38822 | $13 / 01 / 80$ $13 / 01 / 80$ | $66^{\circ} \mathrm{E}$ | $67^{\circ} \mathrm{S}$ $67^{\circ} \mathrm{S}$ | 9 9 | 12/01/81 | $773^{\circ} \mathrm{E}$ | $67^{\circ} \mathrm{S}$ | USSR | 90 | 44278 | 4/02/82 | 5*W | 67*S | 8 | 19/01/84 | $21^{\circ} \mathrm{E}$ | $68^{\circ} \mathrm{S}$ | USSR |
| $43^{7}{ }^{7}$ | 38822 38849 | $13 / 01 / 80$ $14 / 01 / 80$ | $66^{\circ} \mathrm{E}$ | $67 *$ $66^{\circ} \mathrm{S}$ | 9 9 | 19/01/81 | $77^{\circ} \mathrm{E}$ | $68^{\circ} \mathrm{S}$ | USSR | 91 | 45415 | 3/02/82 | 7*W | $70^{\circ} \mathrm{S}$ | 5 | 27/11/82 | $45^{\circ} \mathrm{E}$ | $64^{\circ} \mathrm{S}$ | USSR |
| $44^{8} 8$ | 38849 38851 | $14 / 01 / 80$ $14 / 01 / 80$ | $66^{\circ} \mathrm{E}$ | 66*S | 9 9 | 27/11/81 | 57 $57^{\circ} \mathrm{E}$ | $61^{\circ} \mathrm{S}$ | USSR | 92 | 45434 | 3/02/82 | $6^{6} \mathrm{~W}$ | 70*s | 9 | 17/01/86 | $1^{\circ} \mathrm{E}$ | $6^{69}$ S | USSR |
| 46 | 35344 | 14/01/80 28/01/80 | $26^{6} 7^{\circ} \mathrm{E}$ | 66's | 8 | 27/11/81 | 57 $19^{\circ} \mathrm{E}$ | 61. 69 | USSR USSR | 93 94 | 43902 | 18/01/83 9/01/84 | $84^{\circ} \mathrm{W}$ | 69 $66^{\circ} \mathrm{S}$ | 9 | 20/02/84 | $161^{\circ} \mathrm{W}$ | $71^{\circ} \mathrm{S}$ | Japan |
| $47^{2}$ | 35367 | 29/01/80 | $26^{\circ} \mathrm{E}$ | $69^{\circ} \mathrm{S}$ | 8 | not | covere |  |  |  | 4 | $9 /$ |  | 66 S | 5 | 27/01/84 | $18^{\circ} \mathrm{E}$ | $68^{\circ} \mathrm{S}$ | USSR |

[^10]

Figs. 13 and 14. Position of marking and of recovery of minke whales marked in Area III that show an eastward movement and in Area IV that show a westward movement. Same-season recoveries are not shown.

Table 5

Area of marking and area of recovery of recovered Southern Hemisphere minke whales. Also shown is the distance travelled in an easterly or westerly direction (units are degrees of longitude; a minus sign indicates direction was westerly), the time elapsed between marking and recovery, and estimated length at marking and at recovery. ${ }^{*}=$ Recorded as two whales when marked

| Whale | Area marked | Area recovered | Distance travelled east/west | Time between marking and recovery |  | Length at marking and recovery (m) |  | Whale | Area marked | Area recovered | Distance travelled east/west | Time between marking and recovery |  | Length at marking and recovery (m) |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  | Days | Years) | M | R |  |  |  |  | Days | Years) | M | R |
| 1 | III | III | -1.4 | 1 | (0) | 8.5 | ? | 48 | III | III | -6.8 | 1,811 | (5) | 8. 2 | ? |
| 2 | IV | IV | -7.2 | 384 | (1) | 8.5,8.5* | 8.7 | 49 | III | III | 27.9 | 1,055 | (3) | 8.2 | 9.2 |
| 3 | IV | IV | 14.7 | 1,481 | (4) | 7.9 | ? | 50 | III | III | 24.0 | 1,789 | (5) | ? | ? |
| 4 | IV | IV | -0.1 | 368 | (1) | 8.2 | ? | 51 | III | III | -5.6 | 2,538 | (7) | 9.1 | 9.3 |
| 5 | IV | IV | 31.1 | 1,817 | (5) | ? | 7.9 | 52 | III | II | -53.8 | 891 | (2) | 8.2 | 9.0 |
| 6 | IV | IV | -1.2 | 734 | (2) | 8.5 | ? | 53 | III | III | 12.0 | 694 | (2) | 7.9 | ? |
| 7 | IV | IV | 0.1 | 18 | (0) | 8.5 | ? | 54 | III | III | 31.6 | 1,039 | (3) | 8.5,8.5* | 8.9 |
| 8 | IV | IV | 17.5 | 1,796 | (5) | 8.5 | 9.1 | 55 | III | IV | 73.7 | 1,405 | (4) | 8.2 | 9.0 |
| 9 | IV | III | -42.0 | 2,926 | (8) | ? | ? | 56 | 1II | III | 44.7 | 307 | (1) | 8.2 | ? |
| 10 | IV | IV | 3.1 | 738 | (2) | 8.2 | ? | 57 | III | III | 8.8 | 713 | (2) | 8.5 | 8.5 |
| 11 | IV | IV | 2.9 | 2,938 | (8) | 7. 0 | ? | 58 | III | IV | 66.7 | 1,851 | (5) | 8.5 | 8.6 |
| 12 | IV | III | -27.6 | 1,414 | (4) | 7.0 | 8.7 | 59 | III | III | 11.6 | 1,800 | (5) | 8.2 | ? |
| 13 | IV | III | -67.6 | 410 | (1) | 7.9 | 8.5 | 60 | III | III | 54.3 | 658 | (2) | 8.5 | ? |
| 14 | IV | III | -11.1 | 705 | (2) | 7.9 | 8.6 | 61 | V | V | -7.5 | 1,850 | (5) | ? | ? |
| 15 | IV | IV | 1.2 | 416 | (1) | 7.9 | ? | 62 | V | V | -11.8 | 1,481 | (4) | 7.5 | 8.8 |
| 16 | IV | IV | -23.9 | 713 | (2) | 8.5 | ? | 63 | V | V | -7.8 | 757 | (2) | 8.5 | 7.9 |
| 17 | IV | IV | -26.8 | 1,057 | (3) | 8.2 | 8.0 | 64 | V | V | -13.0 | 1,106 | (3) | 8.2 | 8.6 |
| 18 | IV | V | 39.0 | 2,570 | (7) | 9.1 | 9.5 | 65 | V | V | -8.1 | 753 | (2) | 8.5 | 8.9 |
| 19 | IV | IV | -21.8 | 707 | (2) | 9.1 | 9.2 | 66 | V | V | 2.5 | 1,110 | (3) | 7. 9 | 8.3 |
| 20 | IV | IV | -21.6 | 707 | (2) | ? | ? | 67 | V | V | 1.4 | 21 | (0) | 8.6 | 8.6 |
| 21 | IV | IV | 7.8 | 1,420 | (4) | ? | 8.7 | 68 | V | IV | -33.4 | 382 | (1) | 8.4 | 9.0 |
| 22 | IV | V | 20.6 | 2,891 | (8) | 7.0,7.6* | ? | 69 | V | IV | -30.2 | 704 | (2) | ? | 9.0 |
| 23 | IV | IV | 2.7 | 1,415 | (4) | 7.6,7.9* | 9.0 | 70 | V | V | -6.5 | 2,193 | (6) | 8.5 | 9.6 |
| 24 | IV | IV | -10.8 | 1,448 | (4) | ? | ? | 71 | V | III | -114.7 | 367 | (1) | 8.6 | ? |
| 25 | IV | IV | -9.9 | 2,878 | (8) | 8.5 | 9.5 | 72 | V | V | -11.1 | 2,205 | (6) | 7.9 | 8.1 |
| 26 | IV | III | -75.5 | 1,395 | (4) | 7.0 | ? | 73 | V | IV | -33.2 | 381 | (1) | 8. 7 | 8.1 |
| 27 | IV | IV | -21.5 | 1,398 | (4) | 8.8 | 8.2 | 74 | V | V | -1.2 | 746 | (2) | 8.5 | ? |
| 28 | IV | IV | -11.9 | 18 | (0) | 9.1 | ? | 75 | V | V | -10.8 | 728 | (2) | ? | 8.8 |
| 29 | IV | IV | -12.2 | 18 | (0) | 7.6 | 9.0 | 76 | V | V | -1.3 | 1,104 | (3) | 8.7 | 8.8 |
| 30 | IV | IV | -11.9 | 18 | (0) | 8.5 | ? | 77 | V | V | -10.0 | 730 | (2) | 8.5 | 7.7 |
| 31 | IV | IV | -31.9 | 2,581 | (7) | ? | ? | 78 | V | V | -2.6 | 744 | (2) | 8.7 | 8.5 |
| 32 | IV | IV | -11.5 | 1,068 | (3) | 8. 8 | 8.8 | 79 | V | V | -3.0 | 741 | (2) | 8.5,9.0* | 9.1 |
| 33 | IV | IV | -8.8 | 2,601 | (7) | 7.3 | 8.6 | 80 | V | V | -10.8 | 722 | (2) | 8.6 | ? |
| 34 | IV | IV | -37.8 | 2,921 | (8) | 8. 8 | ? | 81 | V | V | 6.2 | 2,228 | (6) | 8.6 | 8.9 |
| 35 | IV | IV | -18.3 | 1,445 | (4) | 6.7 | ? | 82 | V | IV | -29.7 | 2,167 | (6) | 7.9 | 8.9 |
| 36 | IV | IV | -39.3 | 325 | (1) | 7.0 | $?$ | 83 | V | V | -5.9 | 18 | (0) | ? | 8.7 |
| 37 | V | IV | -2.4 | 19 | (0) | 8.5 | ? | 84 | V | VI | 27.0 | 1,127 | (3) | 8.2 | 8.4 |
| 38 | III | III | -39.8 | 2,211 | (6) | 8.5 | ? | 85 | V | III | -139.2 | 2, 186 | (6) | ? | 8.9 |
| 39 | III | III | -27.1 | 1,808 | (5) | 7. 9 | ? | 86 | V | VI | 20.1 | 762 | (2) | 8. 0, 7. 0* | 8.8 |
| 40 | III | III | 2.2 | 342 | (1) | 8.2 | 7.8 | 87 | V | VI | 16.6 | 1,798 | (5) | 8.5 | 8.3 |
| 41 | III | IV | 6.9 | 1,882 | (5) | 8.5 | 7.8 | 88 | II | II | 1.4 18.3 | 1,333 | (4) | 7.9 8.5 | 8.2 |
| 42 | III | IV | 6.2 | 365 | (1) | 8. 2 | ? | 89 | II | III | 18.3 | 1,067 | (3) | 8.5 | ? |
| 43 | III | IV | 10.2 | 372 | (1) | 8.5 | ? 6 | 90 | II | III | 26.6 | 714 | (2) | 8.3 | ? |
| 44 | III | III | -9.6 | 683 | (2) | 8. 5 | 8.6 | 91 | II | III | 53.2 | + 297 | (1) | ? 5 | ${ }^{?}$ |
| 45 | III | III | -9.6 | 683 | (2) | 8.5 | ? | 92 | II | III | 8.7 -77 | 1,444 | (4) | 8.5 | 8.8 |
| 46 | III | III | -7.3 | 722 | (2) | 8. 5 | ? | 93 | I | VI | -77.2 | 398 | (1) | 8. 5 | 8.0 |
| 47 | III | III | 3.9 | 698 | (2) | 8.2 | ? | (94 | V1 | III | 167.6 | 18 | (0) | ? | ?) |



Fig. 15. Position of marking and of recovery of minke whales marked in Area IV that show an eastward movement. Same-season recoveries are not shown.


Fig. 16. Position of marking and of recovery of minke whales marked in Area $V$ that show a westward movement. Same-season recoveries are not shown.

We might hypothesise that whales are attached to 'ranges', but that the ranges cross boundaries of management Areas. In this case, we would perhaps expect many whales to enter the Area of marking in each successive season, but also to enter one or more additional Areas, so that some are recovered outside of the Area of marking. In this case we would not expect a trend with time in the proportion of recoveries that are from the Area of marking. This hypothesis is consistent with the result from the above test for trend. A further hypothesis that may be of interest is that some whales show a geographic shift in range following marking. If such a shift were permanent, we are unable to test for it, since its effect would be indistinguishable from that which occurs if ranges overlap with Areas. We could hypothesise that there may be a


Fig. 17. Position of marking and of recovery of minke whales marked in Area $V$ that show an eastward movement. Same-season recoveries are not shown.

Table 6
Area of marking and Area of recovery of recovered whales

| Area marked | Area recovered |  |  |  |  |  | Area marked | Area recovered |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | I | II | III | IV | V | VI |  | I | II | III | IV | V | VI |
| I | 0 | 0 | 0 | 0 | 0 | 1 | IV | 0 | 0 | 5 | 28 | 2 | 0 |
| 11 | 0 | 1 | 4 | 0 | 0 | 0 | V | 0 | 0 | 2 | 5 | 18 | 3 |
| III | 0 | 1 | 18 | 5 | 0 | 0 | VI | 0 | 0 | 0 | 0 | 0 | 0 |

Table 7
Cross classification of recoveries: number of seasons between marking and recovery by whether recovery was in the Area of marking

| No. of seasons between | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | Total |
| :--- | ---: | ---: | ---: | ---: | ---: | ---: | ---: | ---: | ---: | ---: |
| marking and recovery: | 7 | 6 | 20 | 7 | 8 | 7 | 4 | 3 | 3 | 65 |
| Same Area recoveries: | 1 | 8 | 5 | 2 | 4 | 3 | 2 | 1 | 2 | 28 |
| Diff. Area recoveries: | 8 | 14 | 25 | 9 | 12 | 10 | 6 | 4 | 5 | 93 |
| Total: |  |  |  |  |  |  |  |  |  |  |

temporary effect, in that the choice of Area in the following season may be influenced by marking. (Same-season recoveries are not used here, since the whale may already be committed to the Area of marking, or may have insufficient time between marking and recovery to move far). A two by two contingency table test with continuity correction of the hypothesis that no temporary effect exists may be carried out. Recoveries are categorised according to whether they occurred one season after marking or more (ignoring same-season recoveries) and also as to whether they were from the Area of marking or not. Using a continuity correction, we obtain a chi-square value with one degree of freedom of 3.68 ( $\mathrm{p}=0.055$ ). Hence, there is some evidence for such an effect. Table 7 indicates that the discrepancy is in the 'expected' direction; the number of recoveries outside the marking Area exceeds the number within only in the case of recovery one season after marking. However, the result is not quite significant at the $5 \%$ level (two-tailed test), so is not conclusive. (Some statisticians would argue that a
one-tailed test is more appropriate here, in which case, the result becomes significant at the $5 \%$ level).

In subsequent analyses, we must consider which marking records and which recovery records we analyse. We turn first to the problem of recoveries outside the Area of marking. As noted above, movement between Areas cannot be regarded as a diffusion process, since the proportion of marked whales remaining in the Area of marking appears not to decline with time. A better model might be to assume that an unknown proportion of whales that use an Area will be in the Area at any given time. Marked whales that use the Area for only a part of the whaling season are relatively less likely to be recovered in that Area. However, if probability of capture in neighbouring Areas is similar, their overall probability of recovery will be close to that for marked whales that stay in the marking Area. It follows that we should include recoveries from outside the Area in abundance analyses. If we do not, the marked population (and its estimated size) will include whales both in the Area and outside, whereas recoveries will be from the sub-population that is using the Area during the whaling season. Stock size would therefore be overestimated. There are sufficient data to attempt analyses for Areas III, IV and V only. We see from Table 1 that, during 1978/79 to 1986/87, similar numbers of whales have been taken from each of these three Areas, with lower numbers from Areas II and VI. Movement between Areas should therefore not affect analyses for Area IV too badly if recoveries outside the Area are included, but we expect too few Area III recoveries from Area II, and too few Area V recoveries from Area VI, so that some overestimation of abundance should be anticipated for these two Areas. Another option is to combine Areas III, IV and V, and to estimate the total population of the three Areas, as was done by IWC (1984, $\mathrm{pp} .82-4)$. This has similar implications; only movement between Areas II and III and between Areas V and VI is likely to cause serious difficulty.

If permanent emigration is rare, but instead marked whales move in and out of the Area in which they were marked, movement between Areas, and inclusion or exclusion of recoveries outside the Area, will have little impact on survival estimation. If permanent emigration occurs, despite the evidence to the contrary provided by Table 7, then inclusion of recoveries outside the Area will provide an estimate of probability of survival, with some underestimation if probability of capture is lower in neighbouring Areas, whereas exclusion of such recoveries would give an estimate of the combined probability that a whale survives and does not permanently emigrate between successive seasons.

The number of whales marked is not well-defined, since there are several categories of 'mark verdict', reflecting information available on whether the mark was successfully placed (see Best and Butterworth, 1980). Table 8 shows that there is a certain amount of uncertainty in the fate of marks, even when the recorded outcome seemed certain. For example, 10 marks have been recovered that were recorded as misses (excluding whale 94 ; see Table 4). We see from Table 8 that $1.7 \%$ of possible hits have subsequently been recovered, compared with $2.6 \%$ of definite hits. Clearly, many successful hits are classified as 'possible'. Problems are sometimes exacerbated by marks breaking up in flight (Butterworth and Best, 1982). Ohsumi (1985) analyses mark recoveries, and estimates that the true number of hits is 1.3 times the

Table 8

Recorded final verdict on marks fired at Southern Hemisphere minke whales, and numbers from each category subsequently recovered. The number of marks recovered (102) exceeds the number of marked whales recovered (93) as a result of multiple-marking. Final verdict 3
was only used when a whale was known to be multiply-marked.
Final verdict definition: $1=$ Invalid (whale believed fatally injured); 2 $=$ Fate unknown, e.g. after accidental or practice firing or misfires over boat side; $3=$ Multiple tag, if the same whale is believed hit more than once the final verdict in one of these records will be 9 (or possibly 1), and 3 in all the others; $4=$ No verdict; $5=$ Miss; $6=$ Ricochet; $7=$ Protruding hit; $8=$ Possible hit; $9=$ Hit (except for multiple tags-see final verdict $=3$ )

| Final verdict | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| :--- | ---: | ---: | ---: | ---: | ---: | ---: | ---: | ---: | ---: |
| Frequency | 28 | 22 | 105 | 104 | 1,646 | 172 | 112 | 1,041 | 2,806 |
| No. recoveries | 0 | 0 | 0 | 1 | 10 | 0 | 1 | 18 | 72 |
| \% recovered | 0.0 | 0.0 | 0.0 | 1.0 | 0.6 | 0.0 | 0.9 | 1.7 | 2.6 |

recorded number of definite hits for non-streamer marks. Joyce (1984; 1985) considers the results of video experiments. Although there were problems in discerning the fate of the mark in a number of video recordings, he estimates that for streamer marks there is $10 \%$ underestimation of the number of hits if definite hits only are used; in the case of non-streamer marks the value is $20-21 \%$. Table 4 shows that, of 93 marked whales recovered, 70 were recorded as definite hits. In terms of marks recovered, 72 of 102 were recorded as definite hits (Table 8). Early mark-recapture analyses of these data are not always explicit in the assumptions made, but it seems that the number effectively marked was often taken to be the number of definite marks (i.e. 'hits') plus those from a 'non-hit' category (which includes 'possible hit' and 'hit protruding') that were subsequently recovered. For example, Brown and Wada (1982) use this strategy. This procedure is clearly biased unless a correction along the lines of Ohsumi (1985) is applied. The compromise adopted in 1983 (IWC, 1984, pp.82-4), in which recoveries from non-hit categories were given a weight of one half, seems arbitrary. We consider the correction of Ohsumi (1985) valid, but note that it adds variability to the analyses (IWC, 1985, p.79); we may achieve a similar outcome more simply by using only those recoveries that were classified as 'hits'. This reduces the number of recovered whales from 93 to 70 , but the alternative of utilising them, in conjunction with an appropriate correction, will be comparable in its effect on precision. In practice, there is some gain in using Ohsumi's approach, since corrections can be estimated from pooled data, and applied to separate Area analyses. The corresponding disadvantages are that we need separate corrections for streamer and non-streamer marks, and we must assume that the corrections are independent of other factors, such as season and vessel. We use here therefore the simpler approach, although this leads to some loss of precision.

If we ignore recoveries from 'non-hit' categories, we assume that all 'hits' were successfully placed. Ohsumi (1985) considers this assumption reasonable. In fact, the difficulty is confounded with the possible problems of short-term mark-shedding and marking mortality, and we return to it later.

The treatment of recoveries from 'non-hit' categories is important for estimation of abundance. Inclusion of them, without a corresponding adjustment to number effectively marked, will reduce the estimated stock size by a factor of roughly $70 / 93=0.75$. However, their exclusion may lead to
overestimation; for example, if $10 \%$ of whales classified as hits were not in fact successfully marked, our estimates will be too high by a factor of roughly 1.11 . Although the two biases act in opposite directions, there is no reason to suppose they will tend to cancel each other out, since they may be of a different order of magnitude. Since one of the 'non-hit' categories is 'possible hit', it seems reasonable to assume that recording a successful hit as a 'non-hit' is more common than recording a miss as a 'hit'. We could therefore carry out two sets of analyses, one excluding 'non-hit' recoveries, and one including them, but both excluding 'non-hit' marks not recovered, since they might represent two extremes. For survival estimation, where ratios of abundance estimates on the marked sub-population are taken, the inclusion or exclusion of 'non-hit' categories is of less consequence, and any bias from this source should be minimal.

## The Petersen and Chapman's modified estimators

In view of the small number of recoveries in any given season, we do not consider the unadjusted Petersen estimator here. Further, because marking sometimes overlapped with the whaling operation, and the assumptions required for the analysis of same-season recoveries seem dubious, we exclude whales recovered in the season of marking from the analyses. This necessitates the replacement of the equation $\mathrm{M}_{\mathrm{i}}=\left(\mathrm{M}_{\mathrm{i}-1}-\mathrm{m}_{\mathrm{i}-1}\right) \mathrm{e}^{-\mu}+\mathrm{R}_{\mathrm{i}} . \mathrm{S}$ (above) by $\mathrm{M}_{\mathrm{i}}=\left(\mathrm{M}_{\mathrm{i}-1}-\mathrm{m}_{\mathrm{i}-1}+\mathrm{R}_{\mathrm{i}-1} . \mathrm{s}\right) \mathrm{e}^{-\mu}$, with notation as before. The same substitution is necessary for other models in which the size of the marked sub-population is estimated in this way.

Table 9
Estimated stock sizes (standard errors in parentheses) under Chapman's modification of the Petersen estimator

| Season | Area |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 111 | IV | v | $\mathrm{III}+\mathrm{IV}+\mathrm{V}$ | ALL |
| 'Non-hit' recoveries not used |  |  |  |  |  |
| 1979/80 |  | $202,000$ $(76,000)$ |  | $\begin{gathered} 649,000 \\ (244,000) \end{gathered}$ | $\begin{gathered} 866,000 \\ (326,000) \end{gathered}$ |
| 1980/81 | $\begin{aligned} & 159,000 \\ & (65,000) \end{aligned}$ | $\begin{aligned} & 235,000 \\ & (88,000) \end{aligned}$ |  | $\begin{gathered} 604,000 \\ (181,000) \end{gathered}$ | $\begin{gathered} 885,000 \\ (266,000) \end{gathered}$ |
| 1981/82 | $\begin{aligned} & 181,000 \\ & (64,000) \end{aligned}$ | $\begin{array}{r} 290,000 \\ (145,000) \end{array}$ | $\begin{gathered} 229,000 \\ (132,000) \end{gathered}$ | $\begin{gathered} 753,000 \\ (226,000) \end{gathered}$ | $\begin{aligned} & 1230,000 \\ & (369,000) \end{aligned}$ |
| 1982/83 | $\begin{aligned} & 147,000 \\ & (65,000) \end{aligned}$ | $\begin{aligned} & 192,000 \\ & (78,000) \end{aligned}$ | $\begin{gathered} 84,000 \\ (28,000) \end{gathered}$ | $\begin{gathered} 456,000 \\ (113,000) \end{gathered}$ | $\begin{gathered} 898,000 \\ (223,000) \end{gathered}$ |
| 1983/84 | $\begin{gathered} 341,000 \\ (196,000) \end{gathered}$ | $\begin{gathered} 464,000 \\ (267,000) \end{gathered}$ | $\begin{gathered} 92,000 \\ (37,000) \end{gathered}$ | $\begin{gathered} 881,000 \\ (310,000) \end{gathered}$ | $\begin{aligned} & 1583,000 \\ & (526,000) \end{aligned}$ |
| 1984/85 | $\begin{aligned} & 155,000 \\ & (69,000) \end{aligned}$ |  | $\begin{aligned} & 131,000 \\ & (75,000) \end{aligned}$ | $\begin{gathered} 863,000 \\ (351,000) \end{gathered}$ | $\begin{aligned} & 1715,000 \\ & (647,000) \end{aligned}$ |
| 1985/86 | $\begin{gathered} 241,000 \\ (139,000) \end{gathered}$ | $\begin{aligned} & 173,000 \\ & (86.000) \end{aligned}$ | $\begin{aligned} & 133,000 \\ & (76,000) \end{aligned}$ | $\begin{gathered} 748,000 \\ (304,000) \end{gathered}$ | $\begin{aligned} & 1340,000 \\ & (472,000) \end{aligned}$ |
| 1986/87 | $\begin{gathered} 191,000 \\ (110,000) \end{gathered}$ | $\begin{gathered} 98,000 \\ (40,000) \end{gathered}$ | $\begin{gathered} 48,000 \\ (19,000) \end{gathered}$ | $\begin{aligned} & 331,000 \\ & (99,000) \end{aligned}$ | $\begin{gathered} 762,000 \\ (229,000) \end{gathered}$ |
| Average | $\begin{aligned} & 202,000 \\ & (42,000) \\ & (26,000) \end{aligned}$ | 236,000 $(50,000)$ $(44,000)$ | $\begin{aligned} & 120,000 \\ & (30,000) \\ & (25,000) \end{aligned}$ | $\begin{aligned} & 661,000 \\ & (86,000) \\ & (68,000) \end{aligned}$ | $\begin{aligned} & 1060,000 \\ & (144,000) \\ & (128,000) \end{aligned}$ |
| 'Non-hit' recoveries used |  |  |  |  |  |
| 1979/80 |  | $\begin{aligned} & 205,000 \\ & (77,000) \end{aligned}$ |  | $\begin{gathered} 656,000 \\ (247,000) \end{gathered}$ | $\begin{gathered} 876,000 \\ (329,000) \end{gathered}$ |
| 1980/81 | $\begin{aligned} & 160,000 \\ & (65,000) \end{aligned}$ | $\begin{aligned} & 204,000 \\ & (72,000) \end{aligned}$ |  | $\begin{gathered} 554,000 \\ (159,000) \end{gathered}$ | $\begin{gathered} 812,000 \\ (233,000) \end{gathered}$ |
| 1981/82 | $\begin{aligned} & 160,000 \\ & (53,000) \end{aligned}$ | $\begin{gathered} 293,000 \\ (146,000) \end{gathered}$ | $\begin{aligned} & 117,000 \\ & (52,000) \end{aligned}$ | $\begin{gathered} 585,000 \\ (155,000) \end{gathered}$ | $\begin{gathered} 955,000 \\ (254,000) \end{gathered}$ |
| 1982/83 | $\begin{aligned} & 148,000 \\ & (66,000) \end{aligned}$ | $\begin{aligned} & 108,000 \\ & (34,000) \end{aligned}$ | $\begin{gathered} 62,000 \\ (18,000) \end{gathered}$ | $\begin{aligned} & 314,000 \\ & (65,000) \end{aligned}$ | $\begin{gathered} 590,000 \\ (119,000) \end{gathered}$ |
| 1983/84 | $\begin{gathered} 342,000 \\ (197,000) \end{gathered}$ | $\begin{gathered} 310,000 \\ (154,000) \end{gathered}$ | $\begin{gathered} 92,000 \\ (37,000) \end{gathered}$ | $\begin{gathered} 773,000 \\ (257,000) \end{gathered}$ | $\begin{aligned} & 1270,000 \\ & (382,000) \end{aligned}$ |
| 1984/85 | $\begin{gathered} 89,000 \\ (31,000) \end{gathered}$ | - | $\begin{aligned} & 131,000 \\ & (75,000) \end{aligned}$ | $\begin{gathered} 541,000 \\ (179,000) \end{gathered}$ | $\begin{aligned} & 1145,000 \\ & (361,000) \end{aligned}$ |
| 1985/86 | $\begin{gathered} 240,000 \\ (138,000) \end{gathered}$ | $\begin{aligned} & 130,000 \\ & (58,000) \end{aligned}$ | $\begin{gathered} 89,000 \\ (44,000) \end{gathered}$ | $\begin{gathered} 534,000 \\ (188,000) \end{gathered}$ | $\begin{aligned} & 1042,000 \\ & (328,000) \end{aligned}$ |
| 1986/87 | $\begin{gathered} 190,000 \\ (109,000) \end{gathered}$ | $\begin{gathered} 82,000 \\ (30,000) \end{gathered}$ | $\begin{gathered} 40,000 \\ (15,000) \end{gathered}$ | $\begin{aligned} & 275,000 \\ & (76,000) \end{aligned}$ | $\begin{gathered} 634,000 \\ (175,000) \end{gathered}$ |
| Average | 190,000 | 190,000 | 88,000 | 529,000 | 915,000 |

Serious difficulties remain, since the number of recoveries in a single season and within one Area is often zero or one. It is not valid to analyse only those cases that have more than a predetermined number of recoveries, since this is in effect equivalent to accepting the small estimates of population size and rejecting the large ones. To avoid potentially serious underestimation of stock sizes, we give individual Area estimates only for Areas III, IV and $V$ in Table 9, and use every estimate for which there was at least one recovery. Also given in Table 9 is the average of the estimates for each Area, with the upper standard error calculated by the first of the methods described by Seber (1982, p.138), and the lower by the second. Chapman's modified estimator will provide an estimate and standard error when there are no recoveries. The estimate clearly has little validity! We therefore ignore the data for Area IV, season 1984/85, for which there were no recoveries, and accept that this may generate a slight downward component to the bias in the average estimate for Area IV. We repeat the estimation procedures for the combined Area III $+\mathrm{IV}+\mathrm{V}$, and for all Areas combined, although this last case should be viewed with suspicion, given both the lower effort and the fewer marks placed in Areas I, II and VI.

## Chapman's multiple sample estimator

In Table 10, we give the estimates generated from Chapman's multiple sample estimator, implemented as described earlier. Estimates are very similar to the average estimates of Table 9. Whether this indicates that the previous method is not seriously influenced by very small numbers of recoveries or that both methods are subject to similar biases is unclear; probably the truth lies somewhere between the two. Sources of bias that are common to all the methods of estimating abundance considered here will be discussed later.

Table 10
Estimated stock size under Chapman's multiple sample estimator

|  | Area | Estimated <br> stock <br> size | Standard <br> error | $95 \%$ confidence <br> interval |  |
| :--- | :---: | ---: | ---: | ---: | ---: |
|  | III | 240,000 | 54,000 | $(163,000-$ | $401,000)$ |
| 'Non-hit' | IV | 291,000 | 59,000 | $(204,000-$ | $460,000)$ |
| recoveries | V | 124,000 | 28,000 | $(83,000$ | $210,000)$ |
| not used | III+IV+V | 685,000 | 88,000 | $(543,000$ | $900,000)$ |
|  | ALL | $1,255,000$ | 177,000 | $(972,000$ | $1,701,000)$ |
|  |  |  |  |  |  |
|  | III | 201,000 | 41,000 | $(141,000$ | $318,000)$ |
| 'Non-hit' | IV | 219,000 | 39,000 | $(161,000-$ | $324,000)$ |
| recoveries | V | 91,000 | 18,000 | $(65,000-$ | $141,000)$ |
| used | III+IV+V | 525,000 | 59,000 | $(428,000-1$ | $665,000)$ |
|  | ALL | 942,000 | 115,000 | $(754,000-1,221,000)$ |  |

Table 11
Estimated stock size, instantaneous death rate and annual probability of survival under a single-release mark-recapture model. Standard errors are given in parentheses

| Area | Estimated stock size | Instantaneous death rate | Prob. of survival | $\mathrm{Chi}^{2}$ (df) |
| :---: | :---: | :---: | :---: | :---: |
| 'Non-hit' recoveries not used |  |  |  |  |
| III | $170,000(76,000)$ | 0.213 (0.126) | 0.808 (0.102) | 1.6 (5) |
| IV | 320,000 ( 134,000 ) | 0.076 (0.094) | 0.927 (0.087) | 7.4 (6) |
| V | 199,000 ( 105,000 ) | -0.043 (0.139) | 1.044 (0.145) | 4.9 (4) |
| 11I+IV+V | 652,000 ( 167,000 ) | 0.103 (0.064) | 0.902 (0.058) | 5.1 (6) |
| 'Non-hit' recoveries used |  |  |  |  |
| III | 173,000 (72,000) | 0.147 (0.111) | 0.863 (0.096) | 6.2 (5) |
| IV | 292,000 (109,000) | 0.026 (0.079) | 0.974 (0.077) | 10.1 (6) |
| V | 119,000 (51,000) | 0.017 (0.119) | 0.983 (0.117) | 5.1 (4) |
| III+IV+V | $559,000(127,000)$ | 0.069 (0.055) | 0.933 (0.051) | 4.2 (6) |

## Single-release methods

The model of Parker (1963), but with the simplifying assumption that the population is stable, was applied to the data from Areas III, IV and V, both including and excluding 'non-hit' recoveries (Table 11). The model provides good fits to the data, but small sample sizes mean that the power of the goodness-of-fit tests is low, and standard errors are high. If we regard the three cohorts of marked whales as a single large cohort, and combine them as if they had all been marked in a single release, we obtain the estimates corresponding to Area III $+\mathrm{IV}+\mathrm{V}$ in Table 11. The results suggest a population of around 600,000 whales in the combined area (SE around 150,000 ), which is very close to the results from the previous two methods. The natural instantaneous death rate is estimated to be close to the previously assumed, but rather arbitrary, value of 0.09 (IWC, 1981, p.111). This suggests that long-term mark-shedding or marking mortality may not be as serious a problem for Southern Hemisphere minke whales as it appears to be for Southern Hemisphere stocks of fin and sei whale (see de la Mare, 1985).

## Recovery models for survival estimation

To apply the models of Brownie et al. (1985), a group of whales that was marked in a single season is regarded as a cohort. Apart from small numbers of whales (from which there have been just two same-season recoveries - see Table 4), there is only a single cohort for each Area. We must therefore combine data across Areas to attempt these analyses. To do this, we assume that both survival rates of minke whales and the probability that a whale is taken are the same in each Area for a given season. Discounting whale 94 (see Table 4), there have been no recoveries from the Area VI cohort, and we do not consider it here. When Areas I and II were included, the models provided unsatisfactory fits to the data. (Chi-square goodness of fit tests yielded $p$ values of 0.07 and 0.003 respectively for models 1 and 2 when 'non-hit' recoveries were excluded from the analyses. The program failed when attempting to fit model 3 to these data). The effort in these Areas was relatively low, and consequently, recovery rates were low. We therefore analyse data from Areas III, IV and V only here.

If we utilise data on whales that were marked in one area and recovered in another, then survival estimation will be unaffected by migration among Areas III, IV and V. Emigration from the combined Area III + IV $+V$ should result in slight underestimation of survival rates, since fishing effort in the remaining area is lower, so that marked whales moving from III $+\mathrm{IV}+\mathrm{V}$ to $\mathrm{I}+\mathrm{II}+\mathrm{VI}$ are less likely to be recovered than whales that stay.

When analyses were carried out on all whales recorded as definite hits plus those that were not definitely marked but were subsequently recovered, poor fits were obtained under each model ( p values for models 1,2 and 3 were $0.006,0.013$ and 0.004 respectively). We therefore exclude recoveries of whales that did not have a final verdict of 9 (definite hit). Inclusion or exclusion of this category of whales is unimportant for estimation of survival, although it has an impact on recovery rate estimation (and, under the modification of the models considered below, abundance estimation).

Goodness of fit tests of models 1,2 and 3 now yield $p$ values of $0.045,0.100$ and 0.153 respectively. These values
are better than above, but still cast some doubt on the adequacy of the models to fit the data. Further tests of model 3 against model $2(p=0.238)$ and model 3 against model $1(\mathrm{p}=0.223)$ suggest that model 3 , which assumes constant survival and recovery rates across time, provides an adequate fit to the data, at least when compared with models 1 and 2 . If we use these recovery models to estimate abundance, we would like to assume that marking takes place immediately before the season starts, and that same-season recoveries are included in the analyses. Because of overlap between whale marking and the whaling season, and because there are doubts about whether random mixing can be assumed for the first season, we do not analyse same-season recoveries here. Under model 3, we can ignore this difficulty initially, since survival estimation and tests for selecting an appropriate model are unaffected. However, estimation of the recovery rate (and abundance estimation, as described below) is affected. We may therefore use the estimated survival rate to adjust the number of whales marked to the number that is expected to survive to the start of the next season. The resulting estimates under model 3 are given in Table 12. In view of the small number of recoveries, the estimates should be treated with caution. The instantaneous death rate for Southern Hemisphere minke whales is currently assumed to be 0.09 (IWC, 1981, p.111), which converts to an annual survival rate of 0.914 . The estimate of Table 12, 0.849 , is an estimate of the annual survival rate in the presence of both natural mortality and fishing mortality. If the effects of the latter are removed, we obtain $\phi=0.855$, with approximate standard error 0.053 and $95 \%$ confidence interval ( $0.752,0.958$ ). The assumed value of 0.914 lies comfortably within this interval.

Table 12
Estimated annual survival rates $\psi$ and $\phi$ (including and excluding fishing mortality), recovery rate and life expectancy of Southern Hemisphere minke whales for the combined Area III $+I V+V$, under Brownie's recovery model with constant survival and recovery rates

| Parameter | Estimate | Standard <br> error | $95 \%$ confidence <br> interval |
| :--- | :---: | :---: | :---: |
| Survival rate, $\boldsymbol{\psi}$ | 0.849 | 0.055 | $(0.742,0.956)$ |
| Survival rate, $\phi$ | 0.855 | 0.052 | $(0.752,0.958)$ |
| Recovery rate | 0.0084 | 0.0017 | $(0.0051,0.0117)$ |
| Life expectancy (yrs) | 6.1 | 2.4 | $(3.4,22.1)$ |

We may take the analyses further here, and calculate Petersen type estimates of abundance. We know the number of whales taken in Area III + IV $+V$ for each year, and the number of these that are marked. In addition, we may use our estimate of survival (0.849) to estimate the instantaneous death rate due to natural and fishing mortality, $v$, and adjust the number of marked whales alive, as described for Chapman's modification of the Petersen estimate above, except that subtraction of the number of marked whales taken in season $i$ is now not required. Hence, we now have $M_{i}=\left(M_{i-1}+R_{i-1} . s\right) e^{-v}$, with $v$ as above, and other notation as before. The resulting abundance and recruitment estimates are given in Table 13. The estimated stock size for Area III + IV $+V$ is slightly lower than under the above models, although the difference is small relative to the precision of the estimates.

Estimated stock size and recruitment, Area III + IV + V, derived from Brownie's recovery model of Table 12

| Season | Estimated <br> stock size | Estimated <br> recruitment | Season | Estimated <br> stock size | Estimated <br> recruitment |
| :--- | :---: | :---: | :---: | :---: | :---: |
| $1979 / 80$ | 715,000 | $-6,000$ | $1985 / 86$ | 614,000 | $-296,000$ |
| $1980 / 81$ | 601,000 | 210,000 | $1986 / 87$ | 225,000 |  |
| $1981 / 82$ | 720,000 | $-218,000$ |  |  |  |
| $1982 / 83$ | 393,000 | 442,000 |  |  |  |
| $1983 / 84$ | 776,000 | 102,000 | Average | 600,000 | 29,000 |
| $1984 / 85$ | 760,000 | $-31,000$ | Std error | 69,000 | 95,000 |

The estimated average recruitment per year, as a proportion of the stock size, is 0.05 . The standard errors of Table 13 are simply the sample standard errors, calculated from the sample standard deviations of the separate season estimates. This ignores correlation between estimates, which is likely to be positive, so that we expect to underestimate true variation.

## Jolly-Seber and related models

The implementation of Jolly-Seber and related models has proven problematic. The principal difficulty arises from the fact that there has only been one season of substantial marking effort in any single Area, whereas the models require annual marking of whales. If a value is assumed for the natural mortality rate, the number of surviving marked whales during any given season may be estimated as for earlier methods. This necessitates additional programming, so that 'off the shelf' software for Jolly-Seber type models is ruled out. For this reason, we analysed data from a single Area using only the standard Jolly-Seber model and the modified Jolly-Seber model of Buckland (1980) from this class of models. However, recoveries were insufficient from any single Area to allow useful estimation. Only when all Areas are combined do we have sufficient data to apply the models sensibly. Unfortunately, another difficulty then arises. For the analysis to be valid, effort should be more or less constant across all Areas. Substantial marking did not occur in Areas I, II and VI until marking in Areas III, IV and V had been completed. Hence, the recoveries, expressed as a percentage of whales marked, declined towards the end of the seasons of marking, since low effort in Areas I, II and VI produced very few recoveries. By contrast, the modified Jolly-Seber model estimates that there was a large influx of new whales (through births or immigration), since this too would lead to a lower recovery rate (the marked whales would be diluted by the new, unmarked whales). The standard model suffers the same problem, but in addition, generates unusable estimates of numbers of 'births', ranging from roughly minus two million to plus infinity! In Table 14, we therefore present results for the combined Area III $+\mathrm{IV}+\mathrm{V}$ only, using the modified model, estimating probability of survival when possible, and assuming an instantaneous death rate of 0.09 , which corresponds to an annual survival rate of 0.914 , otherwise. Standard errors and confidence intervals in Table 14 are obtained by a nonparametric bootstrap method, in which sampling with replacement from the capture histories is carried out. The method is equivalent to a parametric bootstrap, in which deviates are taken from a multinomial distribution fitted to the frequencies of each capture history.

## Table 14

Estimated stock size, probability of survival, number of recruits and probability of capture for the combined Area, III + IV + V, under the modified Jolly-Seber model. Recruited population and stock size in thousands

| Season | Stock size | Prob. of survival | Recruited population | Prob. of capture |
| :---: | :---: | :---: | :---: | :---: |
| 'Non-hit' recoveries not used |  |  |  |  |
| 1978/79 |  | 0.933 (0.164) |  |  |
| 1979/80 | 763 (212) | 0.490 (0.168) | 0 (94) | 0.0078 (0.0042) |
| 1980/81 | 371 (147) | [0.914] ( ) | 90 (104) | 0.0131 (0.0054) |
| 1981/82 | 425 (131) | [0.914] ( - ) | 0 (41) | 0.0117 (0.0029) |
| 1982/83 | 384 (131) | [0.914] ( - ) | 322 (272) | 0.0130 (0.0034) |
| 1983/84 | 668 (304) | [0.914] ( - ) | 76 (349) | 0.0074 (0.0031) |
| 1984/85 | 682 (431) | [0.914] ( - ) | 0 (458) | 0.0056 (0.0021) |
| 1985/86 | 620 (567) |  |  | 0.0059 (0.0020) |
| Average | 559 (182) | 0.676 (0.110) | 81 (99) | 0.0092 (0.0013) |
| 95\% CI | (344, 982) | (0.464, 0.888) | $(35,377)$ |  |
| 'Non-hit' recoveries used |  |  |  |  |
| 1978/79 |  | 1.000 (0.108) |  |  |
| 1979/80 | 721 (140) | 0.479 (0.130) | 0 (45) | 0.0082 (0.0026) |
| 1980/81 | 343 (103) | [0.914] ( - ) | 0 (47) | 0.0142 (0.0049) |
| 1981/82 | 309 (86) | [0.914] ( - ) | 0 (11) | 0.0161 (0.0038) |
| 1982/83 | 278 (79) | [0.914] ( - ) | 253 (133) | 0.0179 (0.0042) |
| 1983/84 | 503 (149) | [0.914] ( - ) | 0 (96) | 0.0099 (0.0030) |
| 1984/85 | 455 (158) | [0.914] ( ) | 0 (369) | 0.0084 (0.0026) |
| 1985/86 | 412 (398) |  |  | 0.0088 (0.0029) |
| Average | 432 (106) | 0.692 (0.088) | 42 (67) | 0.0119 (0.0013) |
| 95\% CI | $(267,653)$ | (0.512, 0.860) | $(6,184)$ |  |

The reduced parameter models of Jolly (1982) were applied to the Area III $+I V+V$ data, and to the data for all Areas. Again, the latter case was found to generate dubious estimates of abundance, and poor fits for reduced parameter models. To apply the models, it was necessary to reduce the data to just one season more than the number of seasons in which significant numbers of marks were placed. This could be achieved in two ways; recoveries more than one season after the last marking cruise could be discarded, or all data after the season of the last marking cruise could be combined as if a single large sample had been taken at the midpoint of the relevant seasons. The former is wasteful of recoveries, of which there are few, so that we choose the latter here. However, this rules out the reduced parameter models with constant probability of capture, since probability of capture for the sample pooled across seasons is very much greater than for the earlier samples. Table 15 shows the results of applying the model that assumes constant probability of survival, but season-dependent probability of capture. Tests of the null hypothesis that probability of survival was constant were

## Table 15

Estimated stock size, probability of survival, number of recruits and probability of capture for the combined Area III + IV +V , under Jolly's reduced parameter model with constant survival probability. Recruited population and stock size in thousands

| Season | Stock size | Prob. of survival | Recruited population | Prob. of capture |
| :---: | :---: | :---: | :---: | :---: |
| 'Non-hit' recoveries not used 1978/79-1980/81 |  | 0.681 (0.121) |  |  |
| 1979/80 | 586 (282) | 0.681 (0.121) | 36 (231) | 0.010 (0.005) |
| 1980/81 | 432 (178) |  | 83 (72) | 0.011 (0.005) |
| 1981/87 | 218 (155) |  |  | 0.119 (0.085) |
| Average | 412 (123) |  | 60 (121) |  |
| Average, 1979-81 | 509 (167) |  |  |  |
| 'Non-hit' recoveries used |  |  |  |  |
| 1978/79-1980/81 |  | 0.715 (0.108) |  |  |
| 1979/80 | 638 (304) |  | -48 (246) | 0.009 (0.004) |
| 1980/81 | 404 (151) |  | 46 (54) | 0.012 (0.004) |
| 1981/87 | 193 (117) |  |  | 0.134 (0.081) |
| Average | 412 (120) |  | -1 (126) |  |
| Average, 1979-81 | 521 (170) |  |  |  |

not significant ( $\chi_{1}^{2}=1.26$ and $1.30 ; p>0.1$ ), suggesting that the assumption is reasonable, although the power of the tests is likely to be low.

Estimates of abundance in Table 15 show an apparent sharp decline over time. This occurs because the estimates of the probability of survival $\phi$ are low. These estimates, and the corresponding very small estimates for $\phi_{2}, \hat{\phi}_{2}=$ 0.490 and $\phi_{2}=0.479$ of Table 14 , should be viewed with suspicion. They seem to arise from a quirk in the data. Although combining all six Areas may lead to substantial bias in abundance estimation, the heterogeneity in effort between Areas is of little significance for estimating survival rates. When data from all six Areas are combined, and we ignore recoveries from the 'non-hit' categories, the standard Jolly-Seber model, which does not constrain survival probability estimates to be less than unity, yields $\phi_{2}=0.543$ (similar to above) and $\phi_{3}=4.823$ ! When these estimates are constrained by the modified method of Buckland (1980), they become $\hat{\phi}_{2}=0.936$ and $\phi_{3}=1.000$, and the average survival rate is estimated by $\dot{\phi}=0.973$, with standard error 0.046 and $95 \%$ confidence interval ( $0.863,1.000$ ). The reduced parameter model with constant yields $\phi=1.004$ with standard error 0.123 . Hence, the full set of data does not contradict the assumption $\phi=0.09$, from which $\phi=0.914$, although the confidence intervals of Table 14 suggest otherwise. The estimates of abundance of Table 15 for 1981/87 in particular should therefore be considered unreliable. The estimated average number of recruits to Area III + IV $+V$ per year, as a proportion of the total population, is 0.15 or 0.10 respectively for the two analyses of Table 14 , and 0.14 or 0.00 for those of Table 15 , although the precision on these estimates is very poor.

It is feasible to implement the models of Jolly (1982) that have constant probability of survival to take full advantage of the minke whale data, without having to either discard recoveries more than one season after the last marking cruise or amalgamate seasons, as done here. We believe that if software was modified for this circumstance, and the model with constant probability of survival and of capture applied, we should obtain to a close approximation the estimates of Tables 12 and 13 , which were generated under the corresponding recovery model of Brownie et al. (1985). (This does not hold for the standard errors of Table 13, which are underestimates). Hence, the estimates of Tables 12 and 13, apart from the standard errors of Table 13, should be considered more reliable than those of Tables 14 and 15 , especially in view of the doubts over the reliability of the survival parameter estimates of Tables 14 and 15.

An important property of Jolly-Seber type models is that if either permanent emigration, or long-term mark-shedding or marking mortality, occurs at a constant rate, abundance estimation is still valid, although survival will be underestimated. Arnason and Mills (1981) show this for the case of mark-shedding. By contrast, methods that rely on an assumed value for the natural mortality rate, including Jolly-Seber type models modified as described above, overestimate abundance in these circumstances. In the presence of short-term mark-shedding or marking mortality, or when some whales recorded as 'hits' are missed, both classes of model overestimate abundance.

## Bias adjustments

Given the uncertainty over the validity of many of the assumptions, it is necessary to quantify as far as possible the effects on the analyses when assumptions fail in
specified ways. Table 16 is an attempt to do this. Note that every non-zero adjustment to abundance estimates in this table is downward, while those for survival estimates are upward. In other words, violations of the assumptions considered here, if they affect estimation at all, lead us to overestimate abundance and underestimate survival rates.

Short-term mark-shedding is potentially a serious problem because we do not have the information to assess whether it occurs commonly or rarely. By contrast, its contribution to bias, if we know the percentage of marks shed shortly after marking, is easily quantified. By 'shortly after marking', we mean before the start of the next whaling season, if we do not use same-season recoveries (as here), and in the absence of long-term mark-shedding. If marks are shed at a constant annual rate, and in addition there is short-term mark-shedding, then 'short-term' refers to those shed between marking and the next whaling season, less the number that would be shed in a single year as a result of long-term shedding.
It has been suggested that if marks lodge in the meat, they are unlikely to be shed, whereas if they lodge in, or protrude into, the blubbel, short-term shedding is likely to

Table 16
Approximate adjustments for abundance and survival estimates of Tables 9-15 given various violations of assumptions. The adjustments apply to those estimates for which 'non-hit' recoveries were not used

| Description of assumption violation | Table number |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\frac{9}{\mathbf{N}}$ | $\frac{10}{\mathrm{~N}}$ | 11 |  | 12, 13, 14 \& 15 |  |  |
|  |  |  | N | $\phi$ | N | $\phi$ | B |
| Short-term markshedding: $x \%$ of marks are shed within one year of marking | -x | -x | -x | 0 | -x | 0 | -x |
| Short-term marking mortality: $x \%$ of whales die within one year of marking | -x | -x | -x | 0 | -x | 0 | -x |
| Long-term markshedding: $x \%$ of marks are shed annually; $\begin{aligned} & y=x \% /(1-\phi), z=x / 7, \\ & w x(5 x+y) / 6 \end{aligned}$ | -x | -x | 0 | $\mathbf{x}$ | $\begin{array}{r} 0 \\ -x \\ -5 z \end{array}$ | x x | $\begin{array}{ll} -y & 1 \\ -x & 2 \\ -w^{3} \end{array}$ |
| Long-term marking mortality: $x \%$ of marked whales die annually; $y=x \% /(1-\phi)$, $z=x / 7, w=(5 x+y) / 6$ | -x | -x | 0 | x | $\begin{array}{r} 0 \\ -x \\ -5 z \end{array}$ | $\mathbf{x}$ - $\mathbf{x}$ | $\begin{aligned} & -y \quad 1 \\ & -x^{2} \\ & -w^{3} \end{aligned}$ |
| Capture prob. outside Area is (100-x)\% of that in Area; $p=p r o p n$ of recoveries outside Area, $q=p /[1-(1-p) x / 100]$ | -qx | -qx | -qx | 0 | -qx | 0 | -qx |
| Combined discovery and reporting rate of tags is ( $100-x$ )\% | -x | -x | -x | 0 | -x | 0 | -x |
| Propn marked during 1978/9 takeable in $1979 / 80$ is $(100-\mathrm{x}) \%$ | $\begin{gathered} -x \\ -x / 8 \end{gathered}$ | -x/8 | -x/8 | 0 | $\begin{gathered} -x \\ -x / m \end{gathered}$ | 0 | $\begin{array}{r} -x^{4} \\ -x / n^{5} \end{array}$ |
| $\mathrm{x} \%$ of 'definite hits' are misses | -x | -x | -x | 0 | -x | 0 | -x |
| $\mathbf{x} \%$ of whales recorded as single-marked are in fact double-marked; $y=100 x /(100+x)$ | -y | -y | -y | 0 | -y | 0 | -y |

${ }^{1}$ Tables 12,13 and 15, and Table 14 up to and including the last release of marked whales.
${ }^{2}$ Estimates after the last release of marked whales, Table 14.
${ }^{3}$ Estimate averaged across seasons, Table 14.
${ }^{+}$Area IV, III + IV + V and 'all-Area' estimates, season 1979/80 only.
${ }^{5}$ Area IV, III + IV + V and 'all-Area' estimates, averaged across seasons.
Table 13: $\mathrm{m}=8$ and $\mathrm{n}=7$. Table 14: $\mathrm{m}=7$ and $\mathrm{n}=6$. Table 15: $\mathrm{m}=3$ and $n=2$.
occur. Of 103 marks recovered; the position in which the mark was lodged was recorded in only 40 cases. None of these were lodged either in the blubber or between the blubber and meat, supporting the above suggestion. Although experimental marks fired into carcases are unlikely to be typical of marks fired into live whales, they are the only source of data for estimating the proportion of marks that lodge in the meat. Kasamatsu, Nishiwaki and Sato (1986) reported that, of 39 marks successfully fired into carcases in the 1984/85 season, 35 lodged in the meat, three were between the blubber and the meat, and one in the blubber. In a similar experiment a year earlier, 36 marks were fired. Of these, one was later found on the deck, although it had been recorded as a hit. Another was found to be protruding from the carcase, four more lodged in the blubber, four between the meat and the blubber, and 26 in the meat (Miyashita and Rowlett, 1985). Kato (1981) reports an experiment carried out in the 1979/80 season, in which 26 of 30 marks were recorded as hits. In fact, two of these had ricocheted, yet in all, 27 were later recovered in tissue: eight in blubber, three between the blubber and meat, and 16 in meat. Of those that lodged in tissue, the proportion of marks that were wholly or partly in blubber was therefore $24 / 101$, suggesting a possible short-term shedding rate of around $24 \%$. However, many of these marks were deliberately fired at unfavourable angles and at large distances, and the two most recent experiments were carried out entirely with streamer marks. The validity of the estimate is therefore uncertain.

Short-term marking mortality seems to be of less concern in that the maximum plausible rate is probably not far from zero, especially given that whales that are seen to be injured or adversely affected after marking are given a final verdict of 'invalid' (see Table 8). Since population size is large relative to sample size, we may regard the effect of say $5 \%$ short-term marking mortality on abundance estimates as being the same as the effect of $5 \%$ short-term mark-shedding. The definition of 'short-term' is the same as for mark-shedding. Best and Butterworth (1980) adopt a value for short-term marking mortality of $5 \%$, estimated from a study by Best (1976) on sperm whales. They consider that the true percentage is likely to be smaller.
Long-term mark-shedding and long-term marking mortality also influence bias in the same manner, if we ignore second-order effects, which are negligible here. Essentially, they lead to underestimation of survival rates but no bias in abundance estimates, if we use a model that provides survival rate estimation, and overestimation of abundance if we assume a specified value for the survival rate that does not take account of long-term mark-shedding or marking mortality. The reason for the relatively complex adjustments for Table 14 estimates is that we must switch from the first category of model to the second after the last 'release' of marked whales. The abundance estimates of Tables 9 and 10, which represent the second category of model, are in good agreement with those of Tables 11 and 13, which represent the first category. Table 14, which is a mixture of the two categories, and Table 15 , which belongs to the first, have lower abundance estimates, which is associated with the low estimate of $\phi_{2}$ (Table 14) or $\phi$ (Table 15), referred to above. Hence, if we accept that this low estimate is an anomaly, there is little evidence for a large contribution to bias from long-term mark-shedding or marking mortality. If we consider the abundance estimates for Area III + IV $+V$ given in Tables $9,10,11$ and 13 , we estimate the
combined annual rate of long-term mark-shedding and marking mortality as $100\{(661+685) /(652+600)-1\}$ $=7.5 \%$. Since abundance estimates are not independent, it is not trivial to obtain a variance for this estimate, and we do not do so here. However, it should be noted that neither of the Area III $+\mathrm{IV}+\mathrm{V}$ survival estimates from Tables 11 and 12 differ significantly from the value assumed for the Table 9 and 10 estimates; indeed, one is almost identical to that assumed value. A similar calculation on the survival estimates of Tables 11 and 12 yields $-100\{(0.855+0.902) /(2 \times 0.914)-1\}=3.9 \%$. Taking the average, we estimate that the combined annual rate for long-term mark-shedding and marking mortality is around $5.7 \%$, although precision is poor, and we cannot rule out the possibility that the true rate is $0 \%$ (or perhaps as high as $15 \%$ ) on this evider ce.
Quantification of the effects of different probabilities of capture for whales within and outside an Area for which abundance estimates are required leads us to rather different conclusions on the most appropriate method to define boundaries to Areas, if new boundaries are to be considered, than the Scientific Committee has reached in the past. If a marked whale moves out of the Area of marking, then for abundance estimation to be unaffected, probability of capture for that whale should remain the same, and recoveries from outside the Area of marking should be included in the analyses. This suggests either that we choose boundaries so that probability of capture is roughly constant either side of each boundary, or that we estimate probability of capture either side of the boundary in some way, and adjust abundance estimates as indicated in Table 16. In 1983 (IWC, 1984, p.87), new Areas were considered, in an attempt to generate boundaries across which whales were less likely to move. Harding criticised this approach (IWC, 1986, p.73), and we too find little to recommend it, given the sparsity of recovery data; further, it seems dubious to us that boundaries exist across which few whales move. Although Area boundaries are arbitrary, we do not believe that it is a worthwhile exercise to experiment with different boundaries. For each of the cruises in which a substantial number of marks was placed, marks were placed throughout the Area with respect to longitude. For the combined Area III $+\mathrm{IV}+\mathrm{V}$, although the whaling fleet of a single nation concentrated within a subset of the region, the Japanese and Soviet whaling fleets, when regarded as a combined operation, generally covered the combined Area relatively uniformly, although in most seasons, there was a gap in the effort in either Area III or Area V or both (Figs 1 to 9 ). Hence, heterogeneity in the operation of the fleets is less than is suggested by examining the Japanese data alone (see for example Shimadzu and Kasamatsu, 1984 and Kasamatsu and Shimadzu, 1986), and any problem arising from heterogeneity is further reduced by the relatively uniform placement of marks. It follows that positioning of boundaries is not critical, unless boundaries exist which few whales cross, which we believe is unlikely. Too few recovery data exist for Areas I, II and VI to allow either reliable estimation or sensible comment on choice of boundaries.

The following approach is a special case of the methods described by Chapman and Junge (1956) and Darroch (1961), and summarised by Seber (1982, pp.431-45), for estimation of abundance and movement. Suppose that we have just two areas, 1 and 2 , with interchange between
them. Suppose further that we have a known number of whales marked in each area, and recoveries from a single season. Let
$\mathrm{N}_{\mathrm{i}}=$ number of whales in area $\mathrm{i}, \mathrm{i}=1,2$
$M_{i}=$ number of whales marked in area $i, i=1,2$
$n_{i}=$ number of whales taken in area $i, i=1,2$
$\mathrm{m}_{\mathrm{ij}}=$ number of whales marked in area i and taken in area $\mathrm{j}, \mathrm{i}, \mathrm{j}=1,2$
$p_{i}=$ probability that a whale in area $i$ is taken, $i=1,2$
$\theta_{i}=$ probability that a whale in area $i$ at marking is also in area $i$ during the whaling season, $i=1,2$
Then $\hat{\mathbf{N}}_{1}=\mathrm{M}_{1}\left(\mathrm{~m}_{22} \mathrm{n}_{1}-\mathrm{m}_{21} \mathrm{n}_{2}\right) /\left(\mathrm{m}_{22} \mathrm{~m}_{11}-\mathrm{m}_{12} \mathrm{~m}_{21}\right)$,

$$
\hat{\mathrm{p}}_{1}=\mathrm{n}_{1} \hat{\mathrm{~N}}_{1} \text { and } \hat{\theta}_{1}=\mathrm{m}_{11} \hat{\mathrm{~N}}_{1} /\left(\mathrm{n}_{1} \mathrm{M}_{1}\right)
$$

Similar estimates follow for area 2. In reality, there are not just two Areas, and the small number of recoveries from Areas I, II and VI rule out a worthwhile analysis involving those Areas. The limitations of the data therefore again force a simplistic application of a potentially useful model. Suppose we define Area IV to be area 1, and Areas III and V together to be area 2. Then we might try the above method to estimate Area IV parameters only, making the assumption that whales that use Area IV do not go as far as Areas II or VI. The method would not be appropriate for area 2 (III +V ) in this case, because we ignore movement between Areas II and III, and Areas V and VI. Another problem now arises, since we need to combine data over seasons, so that sample sizes are adequate. The lack of evidence for a trend in Table 7 suggests that we can assume that the proportion of whales marked in an Area that are also in the same Area during a subsequent season does not vary with time. We might therefore combine data across seasons in a similar manner to that of Sigurjónsson and Gunnlaugsson (1985). If 'non-hit' recoveries are excluded, and seasons 1980/81 to 1986/87 are combined, we obtain $\mathrm{m}_{11}=\Sigma \mathrm{m}_{11 \mathrm{i}}=14, \mathrm{~m}_{12}=4, \mathrm{~m}_{21}=6, \mathrm{~m}_{22}=29, \mathrm{n}_{1} \mathrm{M}_{1}=$ $\Sigma \mathrm{n}_{1 \mathrm{i}} \mathrm{M}_{1 \mathrm{i}}=6.977 \mathrm{x} 10^{6}$, and $\mathrm{n}_{2} \mathrm{M}_{1}=\Sigma \mathrm{n}_{2 \mathrm{i}} \mathrm{M}_{1 \mathrm{i}}=10.737 \times 10^{6}$. Hence, $\hat{\mathbf{N}}_{1}=361,000$ whales in Area IV, and $\hat{\theta}_{1}=0.724$. The estimated stock size is rather higher than earlier estimates, and may indicate that this method has either bias or poor precision (or both). A referee points out that estimation under this model is not robust. Hence, the estimate that roughly $72 \%$ of the whales that use Area IV are in it at any given time during the whaling season should be treated with caution.

It seems reasonable to assume that the discovery rate for marks in whales taken by the Japanese vessels is at or very close to $100 \%$, since electronic scanning equipment is used. Experiments suggest that the combined discovery and reporting rate is also $100 \%$ (Best and Butterworth, 1980; Kato and Miyashita, 1982). However, it is not clear what the corresponding rate for the Soviet operation is. Tillman (IWC, 1982, pp.740-1) and Miyashita (1982a) assume a rate of $70 \%$, and this value is adopted by the Scientific Committee (IWC, 1983, p.95), replacing the previously assumed value of $98 \%$. Tillman (IWC, 1982, p.741) states that 'Best reported...that...experiments revealed a $70 \%$ recovery rate', and Miyashita (1982a) says 'The Soviet figure is based on personal information from the International Observer to the Soviet fleet'. IWC (1981, p.106) reports that 23 of 27 marks inserted on the Soviet factory ship were recovered ( $85 \%$ ), and 15 of 20 inserted 'via catcher boat' were recovered ( $75 \%$ ). However, these results are incomplete, since the 'boilers had not been examined by the time the [Japanese] observer left so that
there might still be more returned marks'. It seems therefore that the estimated discovery and reporting rate for the Soviet operation should be in excess of $80 \%$. A further inconsistency appears to exist; IWC (1981, p.106) reports that ' 17 of 18 marks placed in whales on the Japanese factory ship had been found in the 1979/80 season', suggesting a combined discovery and reporting rate, at least at that time, slightly below $100 \%$. Subsequently, all 18 were reported as having been found (Kato and Miyashita, 1982). There have been 51 marked whales recovered by Japan and 40 by the Soviet Union. When adjusted for the size of the catch in each Area for each season, this indicates that the reporting rate for the Soviet Union is roughly 0.87 times that for Japan, although from this evidence, the difference in reporting rates is not significant ( $\chi_{1}{ }^{2}=0.37 ; p>0.5$ ). In view of the uncertainties, it may not be unreasonable to assume a combined rate, across both fleets, of $90 \%$, which is the value assumed for northeast Atlantic minke whales by Beddington et al. (1984). The resulting bias adjustments can then be made with reference to Table 16, with $x=10 \%$.

Adjustments to estimates of abundance taking into account the proportion of whales takeable appear to have been given little consideration. IWC (1982, p.702) simply assumes that ' $81 \%$ of the marked whales were in the recruited population'. One year later (IWC, 1983, p.95), it was decided that 'the proportion of takeable animals (those over 27 ft ) should be accounted for and varied by Area as seen in sightings cruises'. As noted at the time, small whales grow, and will eventually become takeable. Harding (IWC, 1984, pp.97-8) assumes that small marked whales become takeable one season after marking. Since same-season recoveries are not analysed here, the proportion takeable, as estimated during sightings cruises, is inappropriate. Even if same-season recoveries are used, whales marked are unlikely to be representative of those seen on sightings cruises. Table 3 shows that there is a notable avoidance of small whales in the sample of marked animals in the later years relative to, in particular, the 1978/9 season. Whales are considered 'takeable' if they exceed 8.2 m ( 27 ft ). The average length of whale taken by Japanese pelagic whaling in the Antarctic during 1971/72 to 1979/80 was 27.06 ft for males and 28.10 ft for females (Kato, 1982). Although average length may have increased in recent years, it seems likely that many whales less than 8.2 m are taken, given the difficulty in estimating the lengths of live animals. Hence, in practice, whales slightly shorter than 8.2 m are probably subject to capture. Given that same-season recoveries are not used here, it seems likely that bias as a result of marked whales not being takeable is small, except for the 1979/80 season, which is one season after the marking of substantial numbers of small whales. If we arbitrarily assume that whales less than 7.25 m at marking are not takeable one season later, but become so two seasons after marking, then $x=100 \times 125 /$ $725=17.2 \%$ may be entered in Table 16 to calculate bias adjustments. Corresponding x values for other seasons range from $0 \%$ to $1.5 \%$, and lead to negligible bias adjustments, so are ignored here.

Recording misses as 'definite hits' is probably a small source of error. Such an error is identical in its effect on estimates as short-term mark-shedding.

From Table 4, we can estimate that the percentage of whales recorded as single-marked that were double-marked is $100 \times 6 / 89=6.7 \%$. However, since many marks could not be assigned to a carcase, this may be an
underestimate. Assuming all whales for which there was some doubt were double-marked yields an estimate of $100 \times 10 / 84=11.9 \%$; this is likely to be an overestimate. For the adjustment in Table 16 to be valid, we should use $x=6.7 \%$, since the whales for which there was doubt were all assumed to have been single-marked in the analyses. Christensen and Rørvik (1978 and later) adjusted their estimates in an equivalent way by reducing the effective number marked by the appropriate ratio.

If each assumption violation, $\mathfrak{j}$, could be quantified, and a corresponding bias adjustment, $\mathrm{v}_{\mathrm{j}}$, calculated from Table 16, we could obtain a crude overall adjustment for any given estimate as $\Pi\left(1+v_{j} / 100\right)$. Consider for example the Area III $+\mathrm{IV}+\mathrm{V}$ estimates, excluding 'non-hit' recoveries. First take the average abundance estimate across seasons of Table 9 and the estimate of Table 10 , which have identical adjustments. Suppose the combined effects of short-term mark-shedding (tentatively estimated at $24 \%$ above), short-term marking mortality and recording of misses as hits are a $30 \%$ loss of marks. Then $x_{1}=-v_{1}=30 \%$. Suppose further that long-term mark loss and marking mortality occurs at a rate of $5.7 \%$ per year, as estimated above, so that $x_{2}=-v_{2}=5.7 \%$. During seasons 1979/80 to 1986/87, a total of 16,366 whales was taken from Areas I, II and VI, and 36,661 from Area III $+\mathrm{IV}+\mathrm{V}$. Area III + IV $+V$ spans $190^{\circ}$ of longitude; if we assume it contains 190/170 times as many whales as the remaining Areas, we estimate very approximately that the ratio of probabilities of capture outside and within Area III $+\mathrm{IV}+\mathrm{V}$ is 0.50 , so that $\mathrm{x}_{3}=50 \%$ and $\mathrm{v}_{3}=-4.4 \%$. If the average combined discovery and reporting rate is $10 \%$, then $\mathrm{x}_{4}=-\mathrm{v}_{4}=10 \%$. Assuming that $17.2 \%$ of whales marked in 1978/79 are not takeable during 1979/80 yields $\mathrm{x}_{5}=17.2 \%$ and $\mathrm{v}_{5}=-2.2 \%$. Finally, if $6.7 \%$ of whales recorded as single-marked are double-marked, then $x_{6}=6.7 \%$ and $v_{6}=-6.3 \%$. The overall adjustment is therefore given by $0.70 \times 0.943 \times 0.956 \times 0.90 \times 0.978 \times 0.937=0.520 ; \quad$ the relevant estimates of Tables 9 and 10 should be multiplied by 0.520 under these assumptions, to give an estimated stock size of 344,000 and 357,000 respectively for Area III $+\mathrm{IV}+\mathrm{V}$. The equivalent estimates of abundance from Tables 11 and 13 should be multiplied by 0.552 , that of Table 14 by 0.528 , and that of Table 15 by 0.532 , giving $360,000, \quad 331,000, \quad 296,000$ and 219,000 animals respectively. The final estimate becomes 271,000 if the unreliable 1981/82 estimate is discounted. The corresponding survival estimates of Tables $11,12,14$ and 15 are multiplied by 1.057 , yielding $0.953,0.904$ (adjusted to exclude fishing mortality), 0.715 and 0.720 respectively. The multipliers for recruitment are 0.217 for Table 13, 0.467 for Table 14 and 0.203 for Table 15, so that the estimated average annual recruitment is $6,000,38,000$ and 12,000 animals respectively. (Note the small bias adjustment to the survival estimates, and the very large adjustment to the recruitment estimates of Tables 13 and 15). Estimates may be easily reworked with different values, but valid standard errors or confidence intervals are not available on these adjusted estimates, since the precision of many of the adjustments cannot be quantified.

## RECOMMENDATIONS FOR FUTURE WORK

Pollock (1987) recommended that marking should be carried out either just before or just after the whaling season; if this is not feasible, and marking is carried out during the whaling season, he says that the number of
whales marked should be adjusted to represent the corresponding number still alive just after the whaling season. We concur with this, although in our view, the adjustment should simply be to exclude whales recovered during the same season in which they were marked. Adjustment for natural mortality between marking and the end of the season is inappropriate if the average date of marking is closer to the average date that whales are taken than is the end of the season. If marking is carried out before the season starts, same-season recoveries may be useful. However, in view of the unresolved difficulties connected with non-random mixing of whales and non-random search of whaling vessels, and given the very small number of same-season recoveries, we consider that analyses that use same-season recoveries are of limited value.

Pollock also recommends that deliberate doublemarking of large numbers of whales be considered. Such an exercise has the potential of providing information on the combined probability that a whale is effectively marked and that the mark is not lost shortly thereafter. Hence, all whales classified as a definite or possible hit could be incorporated in the analyses, with an appropriate adjustment for those that were not, in fact, marked, or shed their marks shortly after marking. Further, long-term mark-shedding could be examined by looking for a declining trend with time in the proportion of double-marked whales that are recovered with two marks. However, the scale of the double-marking programme would have to be large. Of 2,804 minke whales recorded as definitely hit from 1975 to 1984 only 70 have been recovered to date. Suppose we assume that all 2,804 whales had been double-marked. Suppose further that there is no long-term mark-shedding and $5 \%$ of marks are shed shortly after marking. Then even under these idealised conditions, without the complication of 'possible hits', the expectation of the standard error on the estimated proportion of marks shed would be about 0.02 (2\%). If many of the hits were only 'possible', so that the combined probability that the mark was successfully placed and was not shed was 0.8 , the estimate of this probability would have a standard error of about 0.04 based on 70 recoveries. If only 500 whales had been double-marked, and $500 \times 70 / 2,804=12$ recovered, the equivalent standard errors are 0.05 and 0.1 , a level of precision which seems unacceptable. Standard errors for other sample sizes may be calculated by noting that the number of double-marked whales recovered with both marks intact will have a binomial distribution with parameters $n=n u m b e r$ of double marked whales recovered with at least one mark intact and $\mathrm{p}^{\prime}=(1-\hat{\mathrm{p}}) /(1+\mathrm{p})$, where p is the probability that a mark is shed (or is not successfully placed). Then, approximately, $S E(p)=\vee\left\{p(1-p)(1+p)^{2} /(2 n)\right\}$. Further discussion on the use of double-marking data is given by Best (1977).

Another recommendation of Pollock (1987) is that further special studies to estimate the probability of detection and reporting of marks should be carried out, and efforts should be made to increase reporting rates. If further extensive marking is to be carried out, we concur with the opinion that more studies are needed, especially in view of the apparently incomplete information available from previous studies. However, it is unclear to us how satisfactory such studies are. Marks placed by researchers on carcases may be more detectable than marks that have been embedded within a whale for some months or years.

Further, the presence of researchers around the carcases, and the possibility that a whaler may suspect that the mark was placed in the whale after it was taken, may affect the reporting rate. Best and Butterworth (1980) give some discussion of these points, and careful consideration should be given to them if future studies of this type are planned.

At the 1987 meeting of the Scientific Committee (IWC, 1988, p.138), a method was discussed by which whales could be marked using 'DNA fingerprinting' (Hoelzel and Dover, 1989). A biopsy dart is fired at an animal, and a small sample of skin is taken. The fingerprinting procedure then makes the whale individually identifiable. If this process proves effective, it should overcome a number of difficulties: long- and short-term mark-shedding and marking mortality; discovery and reporting rates less than $100 \%$; proportion of marked whales takeable unknown; whales recorded as definite hits that are missed; whales recorded as possible hits or misses that are hits; whales recorded as single-marked that are double-marked. As well as the promise of solving all these difficulties, if the marking programme was sufficiently intensive and over many seasons, multiple 'recaptures' of the same animal would allow more rigorous testing of the assumptions of the method than is possible for the recovery data available now. Whales taken by the whaling fleets should also be incorporated in the sampling programme, so that 'recaptures' would be of two kinds: recoveries, which are killed on capture, and 'retraps', which would be whales that are sampled in the marking programme that have previously been sampled at least once. Consideration should be given as to how to combine the two kinds of recapture. If marking took place outside the whaling season, the method of Buckland (1980) could be applied, where recoveries of 'marked' whales are recorded as known deaths between samples. This has two disadvantages. Firstly, known deaths of unmarked whales are not utilised, so that the analysis would not be fully efficient, and secondly, if marking takes place outside the whaling season, it is more difficult to be sure that the population being marked is the population that is subject to capture during the whaling season. A better approach might be to mark during the whaling season, and discard same-season recoveries, as above. Whales taken will then count towards the sample of animals caught, but will be 'losses on capture', whereas the number of whales successfully darted would be the number of live 'releases'. For the method to be successful, marking should be continued within one Area for at least two successive seasons, to allow the reduced parameter models of Jolly (1982) to be applied, as recommended by Pollock (1987). In practice, a minimum of three seasons should be considered, with of the order of 500 whales or more marked per season. If fewer were marked, recoveries would be insufficient, as is shown by the relatively few recoveries analysed here. Indeed, if harvesting takes place at a lower level in future seasons relative to the past, it would be advisable to mark larger numbers, to increase both the expected number of recoveries and the expected number of 'retraps' - animals sampled during the marking operation that had been sampled in a previous season. The major assumption not directly addressed by the above approach relates to movement of whales between Areas. However, animals taken by the whaling fleets in other Areas will provide information on recoveries of 'marked' whales outside the Area of marking, so that bias adjustments for movement may be estimated. If marking is
also carried out in other Areas, the method described above for estimating numbers moving between Areas may prove useful. De la Mare and Payne (1988) note that multiple recaptures of a single animal will in itself facilitate studies of migration and range.

Photo-identification techniques provide an alternative method of 'marking' minke whales. Opinion seems to be divided on the practicality of the method. If it proves easier to photograph a minke whale such that it is individually identifiable than to obtain a biopsy sample, then serious consideration should be given to the method. However, the fingerprinting method may still prove to be superior. For example, de la Mare and Payne (1988) note that a DNA fingerprint cannot change over time, whereas natural markings may. Further, if some minke whales are not readily identified from a photograph, because distinctive markings are not evident, whereas others are easily identified, there will be a source of heterogeneity in the capture probabilities not present for the fingerprinting method. For this reason, if photo-identification is to be attempted, we recommend that matching of photographs is done by computer, using a method similar to that used by Hiby and Lovell (in press) for grey seals.

If a programme of the above type, using either DNA fingerprinting or photo-identification, is seriously contemplated, consideration should be given to sample size. As shown by Cooke (1986), mark-recapture is not a practical method for estimating trends in stock size. Although his analyses could be repeated with different assumptions, the conclusions are unlikely to be very much less pessimistic than his. As he notes, 'mark-recapture methods do not provide an adequate feedback mechanism with which to regulate a whale fishery...'. It follows therefore that we should not take mark-recapture estimates in isolation, but we should use them in conjunction with results from management feedback techniques (as considered for example by de la Mare, 1986), line transect sampling and CPUE methods. This might be done formally using Bayesian methods, but such an approach is likely to require considerable methodological development, and a less formal amalgamation of information from the various sources is suggested.

## DISCUSSION

The statement by Chapman, de la Mare, Holt and Van Beek (1982) emphasised the uncertainty surrounding stock size estimates. Since it was written, considerably more data have accrued from which to estimate abundance of Southern Hemisphere minke whale stocks. However, it is clear that many uncertainties still exist. Given the size of the whale stocks, and of the region they inhabit, a considerable expenditure of effort yields relatively little reward. The mark-recapture experiments reflect this. The large marking programmes have, between them, given just 93 recoveries; very few data from which to estimate many possible parameters of interest. There are insufficient data to attempt sensible mark-recapture analysis of stock size for Areas I, II and VI. In Table 18, we show estimates of abundance for Areas III, IV and V, and the combined Area III + IV +V, for which bias adjustments have been applied as described above. The ratio of probability of capture outside to within each individual Area was calculated as described for Area III+IV+V above, but using only neighbouring Areas. For example, the Area III

## Table 17

Values assigned to the adjustments of Table 16 to generate the estimates of Table 18

| Description of assumption violation | Value chosen for x | Source |
| :---: | :---: | :---: |
| Short-term markshedding: $\pi \%$ of marks are shed within one year of marking | 24\% | Estimated from proportion of marks that lodged wholly or partly in blubber when fired at carcases ( $24 / 101$ ) |
| Short-term marking mortality: $x \%$ of whales die within one year of marking | 5\% | Best and Butterworth (1980) |
| Long-term markshedding and/or marking mortality: $\mathrm{x} \%$ of marked whales either die or shed their marks annually | 5.7\% | Estimated from discrepancies between estimates obtained assuming a fixed mortality rate and those obtained after estimating survival rates from the data |
| Capture prob. outside Area is $(100-x) \%$ of that in Area | 50\% | Estimated from observed marking and recovery locations and observed catches |
| Combined discovery and reporting rate of tags is $(100-\mathrm{x}) \%$ | 10\% | Estimated from carcase experiments |
| Propn marked during 1978/9 takeable in $1979 / 80$ is $(100-x) \%$ | 17.2\% | Estimated from the observed length distribution of whales marked during 1978/9 |
| $x \%$ of 'definite hits' are misses |  | Assumed small; chosen such that combined effects of short-term mark-shedding (24\%), short-term marking mortality ( $5 \%$ ) and this yields $x=30 \%$ |
| x\% of whales recorded as single-marked are in fact double-marked | 6.7\% | Estimated from recoveries of whales recorded as singlemarked |

Table 18
Estimates of abundance from mark-recapture and line transect analyses of Southern Hemisphere minke whale data from Areas III, IV, V and the combined Area III + IV +V . The mark-recapture estimates have been adjusted for bias, although their standard errors ignore the component of variance due to estimation of the bias adjustments

| Area | III | IV | V | III + IV+V |
| :---: | :---: | :---: | :---: | :---: |
| Average of modified Petersen estimates | $\begin{aligned} & 112,000 \\ & (23,000) \end{aligned}$ | $\begin{aligned} & 128,000 \\ & (27,000) \end{aligned}$ | $\begin{gathered} 67,000 \\ (16,000) \end{gathered}$ | $\begin{aligned} & 345,000 \\ & (45,000) \end{aligned}$ |
| Chapman's multiple sample estimate | $\begin{aligned} & 133,000 \\ & (30,000) \end{aligned}$ | $\begin{aligned} & 158,000 \\ & (32,000) \end{aligned}$ | $\begin{gathered} 69,000 \\ (16,000) \end{gathered}$ | $\begin{aligned} & 356,000 \\ & (46,000) \end{aligned}$ |
| Parker's stable population estimate | $\begin{aligned} & 100,000 \\ & (45,000) \end{aligned}$ | $\begin{aligned} & 184,000 \\ & (77,000) \end{aligned}$ | $\begin{aligned} & 117,000 \\ & (61,000) \end{aligned}$ | $\begin{aligned} & 360,000 \\ & (93,000) \end{aligned}$ |
| Brownie's recovery model estimate |  |  |  | $\begin{aligned} & 331,000 \\ & (38,000) \end{aligned}$ |
| Modified JollySeber estimate |  |  |  | $\begin{aligned} & 296,000 \\ & (96,000) \end{aligned}$ |
| Constant $\phi$ Jolly estimate |  |  |  | $\begin{aligned} & 219,000 \\ & (65,000) \end{aligned}$ |
| Constant $\phi$ Jolly estimate, excluding 1981/87 estimate |  |  |  | $\begin{aligned} & 271,000 \\ & (89,000) \end{aligned}$ |
| Line transect estimates, takeable only | $\begin{gathered} 69,000 \\ (13,000) \end{gathered}$ | $\begin{gathered} 60,000 \\ (10,000) \end{gathered}$ | $\begin{aligned} & 106,000 \\ & (22,000) \end{aligned}$ | $\begin{aligned} & 236,000 \\ & (28,000) \end{aligned}$ |
| Line transect estimates, all animals | $\begin{aligned} & 105,000 \\ & (19,000) \end{aligned}$ | $\begin{gathered} 91,000 \\ (14,000) \end{gathered}$ | $\begin{aligned} & 161,000 \\ & (34,000) \end{aligned}$ | $\begin{aligned} & 358,000 \\ & (41,000) \end{aligned}$ |

ratio was estimated by considering number of whales taken in Areas II and IV relative to the number in Area III, and adjusting for relative size of the Areas. Hence, the estimated ratios are 0.99 for Area III, 0.72 for Area IV and 0.99 for Area V. The standard errors of Table 18 ignore the component of variance from estimation (or 'guesstimation') of the adjustments, and should therefore be regarded as lower bounds to the true standard errors.

Also shown in Table 18 are the corresponding estimates of abundance from analyses of the line transect data from IWC/IDCR cruises, taken from Butterworth, Buckland, Kishino and Silberbauer (1987). The relative consistency between different mark-recapture estimates is not in reality much cause for comfort; they all utilise the same data, and as shown in Table 16, are all subject to much the same biases. The line transect estimates of abundance for the takeable population are on average lower. Several reasons can be suggested for this. Butterworth et al. assumed that the proportion takeable was 0.658 , so that all estimates were multiplied by this figure. Mark-recapture estimates cannot be adjusted in such a simplistic manner. It is likely that whales just short of 27 ft , which in theory are not takeable, have a smaller, but non-zero, probability of capture relative to larger whales (see above). In effect, therefore, the proportion of the population that is at risk of capture might be considerably higher than 0.658 , leading to higher mark-recapture estimates relative to line transect estimates. We thus show line transect estimates for the whole population in addition to those for the takeable population in Table 18. In the absence of other biases, we might expect mark-recapture estimates to lie between the two sets of line transect estimates. For Area III + IV $+V$, five mark-recapture estimates lie within this range, one is smaller and one larger. The individual Area estimates are more variable, with Area III mark-recapture estimates being a little high, Area IV estimates markedly high and Area V estimates low relative to line transect estimates.

The line transect method yields an estimate of the size of the stock in the survey area at the time of the survey. Hence, it excludes whales that are in the pack ice, and whales that are north of the northern limit of the survey. Whales that spend part of the season outside the survey area and part close to the ice-edge are subject to capture, so that the mark-recapture estimates may relate to a larger population of animals than the line transect estimates, as noted by Yamamura (IWC, 1986, p.70).

Another point to consider is that variances for estimates under closed population models are underestimated, since we assume $\mu$ is known to be 0.09 . Further, our variance of the abundance estimate from the recovery model of Brownie et al. (1985) is also an underestimate. If we compare the abundance estimates under the modification of Parker's single release model, under the modified Jolly-Seber model or under the model of Jolly that assumes probability of survival is constant with those from the line transect analyses in Table 18, none of the differences are significant at the $5 \%$ level. Similarly, although mark-recapture methods suggest that the size of the Area IV stock exceeds that of Area V, while line transect estimates indicate the reverse, the apparent discrepancy may mean little when underestimation of variances is taken into account. Possible unquantified biases in the line transect estimates and variances, and unknown movement patterns between Areas, further confuse comparisons between estimates. It has been suggested that pooling across Areas leads to more robust mark-recapture estimates. Such a strategy reduces problems that arise from small sample sizes and from animal movements, both of which may be responsible for the apparent inconsistency between Area IV and Area V mark-recapture and line transect estimates. Thus, comparisons between the Area III + IV +V estimates may be more valid. However, by pooling across Areas, heterogeneity in the probability that a whale is marked or taken becomes more of a concern. A
simulation study may yield insights into these issues, although it might prove difficult to assess the reliability of the conclusions from such an exercise.

Mark-recapture estimates may be higher than those from line transect on average partly because at least one of the assumption violations for the mark-recapture methods is more serious than has been assumed. P.S. Hammond (pers. comm.) considers that short-term mark-shedding might be common, and there is little information in the available data either to support or to refute this possibility. Factors noted above lead us to believe that the population being estimated by mark-recapture methods is wider both geographically and in size distribution than the takeable population as defined for the line transect method. The likely effect of short-term mark-shedding may therefore be less than any discrepancies between the respective mark-recapture and line transect estimates may suggest; indeed, if underestimation of variance is allowed for, the estimated rate for short-term mark loss is well within the range of values for which the mark-recapture and line transect estimates of abundance in Area III $+\mathrm{IV}+\mathrm{V}$ are consistent with each other.

There is little evidence of long-term mark-shedding for Southern Hemisphere minke whales. However, given the low precision on estimates, it is not possible to rule out the existence of a moderate rate of long-term mark-shedding. Nevertheless, the problem would appear to be of smaller magnitude than for Southern Hemisphere stocks of fin and sei whales (de la Mare, 1985). J. Horwood (pers. comm.) suggests an alternative explanation of the sei whale results. Most marks were placed in sei whales early in the period of substantial exploitation. As the fishery developed, the fleet shifted northwards and the size of whales taken declined, so that relatively few recoveries would be expected of older whales that had been marked some years earlier. However, the results from the two fin whale data sets cannot be explained in this way.

Best (IWC, 1985, p.79) considers that high shedding rates are unlikely for Southern Hemisphere minke whales, since 'attempts had been made to disinfect marks before firing. Examination of recoveries ...so far had not indicated any clear signs of rejection.' Although there is a field for reported wound condition in the recoveries data available to us, for minke whales, this was always coded as zero (unknown/not reported). A discussion of the likelihood of abscess formation, and subsequent shedding of the mark, based on comments from three veterinarians, can be found in Best and Butterworth (1980). It was considered that sterilisation of the marks was impractical, and that 'it might be a reasonable (but probably less effective) alternative to attempt to disinfect the marks immediately prior to firing...'.

Christensen and Rørvik (1980) considered the effect of changes in sex ratio in the catch of north-east Atlantic minke whales, and concluded that the mark-recapture estimates, ignoring such changes, were insensitive to them. In the Antarctic, females are more abundant than males close to the ice-edge (Kasamatsu and Ohsumi, 1981), so that their probability of capture is higher. Shimadzu (1982) shows that the ratio is generally close to unity, except for Areas II and III. Although consideration of the ratio is important for the management of the stock, we do not believe that its impact on mark-recapture estimates presented here is large.

Christensen and Rørvik (1981a) also considered the effect of changing availability with age on mark-recapture
estimates, and found that their estimated abundance for northeast Atlantic minke whales roughly doubled. The availability was estimated using the age composition of the stock (Christensen and Rørvik, 1981b). They conclude that, after the age of eight, minke whales become less available to the Norwegian small whale whalers, probably due to increasing caution of animals as they get older. As pointed out by J. Horwood (pers. comm.), whales in the northern fishery are 'stalked', whereas the large catcher boats in the Antarctic chase the whales. Hence, once the whales become takeable, it is unlikely that availability is particularly dependent on age for the Southern Hemisphere stocks.

We exclude same-season recoveries from our analyses. There were only eight of these. Two of the eight come from very small numbers of whales marked ( 24 and 8 respectively; see Table 2), and raise serious doubts about non-random mixing, which was a concern expressed by IWC (1985, p.79). One of these was recovered just one day after marking. Of the remaining six, three were marked on the same day, and 18 days later, recovered on the same day. Only one of the marks could be assigned to carcase, so that the possibility of double- or triple-marking cannot be ruled out. If the marks were placed in different whales, the assumption of independence is clearly invalid. We therefore consider that analysis of same-season recoveries has little value here.

Two aspects of mark-recovery data that might be of interest, but which we have not addressed, relate to information they provide on growth and school structure or integrity. We believe that the available data shed little light on either. From Table 5, there were 46 recoveries for which there is both an estimated length at marking and a measured length at recovery. For ten of these, the length at recovery was shorter than the estimated length at marking. A further five were double-marked, but had been recorded as ten single-marked whales. Only two were same-season recoveries; while for one of these, the estimated and measured lengths were the same, the other was estimated to be 7.6 m , and was measured at 9.0 m eighteen days later. This whale was one of the 'three' noted above, that were marked at the same time and recovered on the same day, and for which the doubt exists as to whether the marks were fired into three whales, two or just one. The other two estimated lengths at marking were 9.1 m and 8.5 m , much closer to the length of the measured whale than the above estimate. Unless age of recovered whales is estimated, the mark-recovery data seem to provide little useful information on growth; even with age estimates, the inaccuracies of estimated lengths at marking may render any analyses on such a small sample meaningless.

The above example of either one, two or three animals marked together and recovered together also illustrates the impossibility of examining school structure and integrity from these data. Whenever two or more animals from a pod were marked, and later recovered together, either they were found to have been the same animal, inadvertantly double-marked, or at most one of the marks could be assigned to a carcase, so that there is doubt whether more than one animal was involved. On no occasion were two marks recovered together that could be assigned to different carcases.

A further aspect of the data that we have not examined in detail is the information they provide on movements. We show the positions of marking and recovery in Figs 11 to 17 , but consider that recoveries are too few to be able to
identify biological stock boundaries. Fig. 14 suggests that there may be a boundary at about $75^{\circ} \mathrm{E}$, close to the Area III/Area IV boundary, with just one whale recorded as crossing it. However, Figs 13 and 15 each show one whale crossing it, and Fig. 16 shows two long distance movements, the shortest route for which would have taken them across $75^{\circ} \mathrm{E}$. Given the small numbers of recoveries, it is inevitable that some lines of longitude will have few examples of recorded whale crossings, especially in the presence of heterogeneity in distribution of effort by the fishery, or in density of whales.

We are also doubtful whether biological boundaries can be adequately represented by a single line of longitude, given that stocks will inevitably mix to some degree, and that any boundary may vary with season, or with oceanographic conditions. If an assessment is to be made of the possible position of biological stock boundaries, we feel that biological information must be utilised in conjunction with the information of Figs 11 to 17.

## ACKNOWLEDGEMENTS

A number of people have been helpful in enabling us to complete this work in the short time available. We would like to thank in particular C. Allison (who supplied the data and Figs 1-9), G. Donovan and R. Gambell. R. Cormack, P. Hammond, J. Horwood, J. Cooke and E. Harding all made time at short notice in heavy schedules, and we are especially grateful for their comments and suggestions. The encouragement of G. Kirkwood was also most welcome, as were the written comments received from D. Butterworth. In addition, we would like to thank our colleagues in the Scottish Agricultural Statistics Service, and in the computing section of the Macaulay Land Use Research Institute, whose willing cooperation eased our task considerably. K. Pollock (North Carolina State University) and M. Nicholson (Fisheries Laboratory, Lowestoft) reviewed the manuscript.

## REFERENCES

Arnason, A.N. and Mills, K.H. 1981. Bias and loss of precision due to tag loss in Jolly-Seber estimates for mark-recapture experiments. Can. J. Fish. Aquat. Sci. 38: 1,077-95.
Bailey, N.T.J. 1951. On estimating the size of mobile populations from capture-recapture data. Biometrika 38: 293-306.
Bailey, N.T.J. 1952. Improvements in the interpretation of recapture data. J. Animal Ecol. 21: 120-7.
Beddington, J.R., Cooke, J.G., Christensen, I., Øritsland, T., Øien, N. and Rørvik, C.J. 1984. Assessments of the northeast Atlantic stock of minke whales. Rep. int. Whal. Commn 34: 285-91.
Best, P.B. 1976. Status of whale stocks off South Africa, 1974. Rep. int. Whal. Commn 26: 264-86.
Best, P.B. 1977. Status of whale stocks off South Africa, 1975. Rep. int. Whal. Commn 27: 116-21.
Best, P.B. and Butterworth, D.S. 1980. Report of the Southern Hemisphere minke whale assessment cruise, 1978/79. Rep. int. Whal. Commn 30: 257-83.
Brown, S.G. and Best, P.B. 1981. Recoveries of whale marks from minke whales in the Southern Hemisphere. Rep. int. Whal. Commn 31: 357-8.
Brown, S.G. and Wada, S. 1982. Southern Hemisphere marking of minke whales. Rep. int. Whal. Commn 32: 803-8.
Brownie, C., Anderson, D.R., Burnham, K.P. and Robson, D.S. 1985. Statistical Inference from Band Recovery Data-a Handbook. 2nd ed. United States Department of the Interior, Resource Publication No. 156.
Buckland, S.T. 1980. A modified analysis of the Jolly-Seber capture-recapture model. Biometrics 36: 419-35.
Butterworth, D.S. and Best, P.B. 1982. Report of the Southern Hemisphere minke whale assessment cruise, 1980/81. Rep. int. Whal. Commn 32: 835-74.

Butterworth, D.S., Buckland, S.T., Kishino, H. and Silberbauer, M.J. 1987. Revised sighting estimates of Antarctic minke whale abundance from the $1978 / 79$ to 1984/85 IWC/IDCR assessment cruises. Paper SC/39/Mil8 presented to the IWC Scientific Committee, June 1987 (unpublished).
Chapman, D.G. 1951. Some properties of the hypergeometric distribution with applications to zoological censuses. Univ. Calif. Public. Stat. 1: 131-60.
Chapman, D.G. 1952. Inverse, multiple and sequential sample censuses. Biometrics 8: 286-306
Chapman, D.G. 1954. The estimation of biological populations. Ann. Math. Statist. 25: 1-15.
Chapman, D.G. 1965. The estimation of mortality and recruitment from a single-tagging experiment. Biometrics 21: 529-42.
Chapman, D.G. 1970. Reanalysis of Antarctic fin whale population data. Rep. int. Whal. Commn 20: 54-9.
Chapman, D.G., de la Mare, W.K., Holt, S.J. and Van Beek, J.G. 1982. Statement on minke whale assessments for the Southern Hemisphere and advice to the Commission. Rep. int. Whal. Commn 32: 136-8.
Chapman, D.G. and Junge, C.O. 1956. The estimation of the size of a stratified animal population. Ann. Math. Statist. 27: 375-89.
Christensen, I. and Rørvik, C.J. 1978. Preliminary results from recent markings of minke whales in the north east Atlantic. Rep. int. Whal. Commn 28: 183-4.
Christensen, I. and Rørvik, C.J. 1979. Stock estimate of minke whales in the Svalbard-Norway-British Isles area from markings and recoveries 1974-77. Rep. int. Whal. Commn 29: 461-2.
Christensen, I. and Rørvik, C.J. 1980. Results from markings of minke whales in the northeast Atlantic. Rep. int. Whal. Commn 30: 201-3.
Christensen, I. and Rørvik, C.J. 1981a. Analysis of markings and recaptures of minke whales in the Barents Sea 1974-79. Rep. int. Whal. Commn 31: 255-7.
Christensen, I. and Rørvik, C.J. 1981b. Availability of minke whales in the Barents Sea and adjacent waters. Rep. int. Whal. Commn 31: 259-62.
Cooke, J.G. 1986. On the utility of mark-recapture experiments for the detection of trends in whale stocks. Rep. int. Whal. Commn 36: 451-5.
Darroch, J.N. 1961. The two-sample capture-recapture census when tagging and sampling are stratified. Biometrika 48: 241-60.
de la Mare, W.K. 1985. Some evidence for mark shedding with discovery whale marks. Rep. int. Whal. Commn 35: 477-86.
de la Mare, W.K. 1986. Simulation studies on management procedures. Rep. int. Whal. Commn 36: 429-50.
de la Mare, W.K. and Payne, R. 1988. The possible use of DNA fingerprinting techniques in making mark-recapture estimates of whale populations. Report of the Scientific Committee, Annex Q. Rep. int. Whal. Commn 38: 138.
Garrod, D.J. and Brown, S.G. 1980. Southern Hemisphere sei whales: a population estimate from marking data. Rep. int. Whal. Commn 30: 507-11.
Hiby, A.R. and Lovell, P. In Press. Computer-aided matching of seal pelage markings. Rep. int. Whal. Commn (special issue 12)
Hoelzel, A.R. and Dover, G.A. 1989. Molecular techniques for examining genetic variation and stock identity in cetacean species. Paper SC/39/O 7 (published in this volume).
Horwood, J.W. 1981. Results from the IWC/IDCR minke marking and sightings cruise, 1979/80. Rep. int. Whal. Commn 31: 287-313.
International Whaling Commission. 1981. Report of the minke whale sub-committee. Rep. int. Whal. Commn 31: 103-21.
International Whaling Commission. 1982. Report of the special meeting on Southern Hemisphere minke whales. Rep. int. Whal. Commn 32: 697-745.
International Whaling Commission. 1983. Report of the sub-committee on minke whales. Rep. int. Whal. Commn 33: 91-122.
International Whaling Commission. 1984. Report of the sub-committee on Southern Hemisphere minke whales. Rep. int. Whal. Commn 34: 77-101.
International Whaling Commission. 1985. Report of the sub-committee on Southern Hemisphere minke whales. Rep. int. Whal. Commn 35: 75-89.
International Whaling Commission. 1986. Report of the sub-committee on Southern Hemisphere minke whales. Rep. int. Whal. Commn 36: 66-78.
Ivashin, M.V. 1981a. USSR. Progress Report on Cetacean Research June 1979-May 1980. Rep. int. Whal. Commn 31: 221-6.

Ivashin, M.V. 1981b. Marking of minke whales (Balaenoptera acutorostrata, Lac. 1804) in the waters of the Southern Hemisphere (Soviet program of tagging). Paper SC/Jn81/MiS33 presented to the IWC Scientific Committee, June 1981 (unpublished).
Jolly, G.M. 1965. Explicit estimates from capture-recapture data with both death and immigration-stochastic model. Biometrika 52: 225-47.
Jolly, G.M. 1982. Mark-recapture models with parameters constant in time. Biometrics 38: 301-21.
Joyce, G.G. 1984. An examination of marking success with . 410 Discovery marks. Rep. int. Whal. Commn 34: 281-3.
Joyce, G.G. 1985. Further observations on marking success with .410 Discovery marks. Rep. int. Whal. Commn 35: 231-3.
Kasamatsu, F., Nishiwaki, S. and Sato, M. 1986. Results of the test firing of improved .410 streamer marks, February 1985. Rep. int. Whal. Commn 36: 201-4.
Kasamatsu, F. and Ohsumi, S. 1981. Distribution pattern of minke whales in the Antarctic with special reference to the sex ratio in the catch. Rep. int. Whal. Commn 31: 345-8.
Kasamatsu, F. and Shimadzu, Y. 1986. Operating pattern of Antarctic minke whaling by the Japanese expedition in the 1984/85 season. Rep. int. Whal. Commn 36: 205-6.
Kato, H. 1981. Trial-firing of .410 marks at a minke whale carcase on the factory ship, 1979/80 Antarctic season. Rep. int. Whal. Commn 31: 367-370.
Kato, H. 1982. Some biological parameters for the Antarctic minke whale. Rep. int. Whal. Commn 32: 935-45.
Kato, H. and Miyashita, T. 1982. Summary of the experiments on .410 mark recovery efficiency on the Japanese factory ship. Rep. int. Whal. Commn 32: 901-3.
Miyashita, T. 1982a. Estimation of the population size of minke whales in Areas III and IV in 1980/81 using a mark-recapture method. Rep. int. Whal. Commn 32: 897-8.
Miyashita, T. 1982b. Estimation of the population size of minke whales in Area I using sightings and marking data. Rep. int. Whal. Commn 32: 911-7.
Miyashita, T. 1983. Estimates of the population size of the Antarctic minke whale using various mark recapture methods. Rep. int. Whal. Commn 33: 379-82.
Miyashita, T. and Kasamatsu, F. 1985. Population assessment of the western North Pacific stock of Bryde's whales. Rep. int. Whal. Commn 35: 363-8.
Miyashita, T. and Rowlett, R.A. 1985. Test-firing of .410 streamer marks. Rep. int. Whal. Commn 35: 305-8.
Ohsumi, S. 1974. Bryde's whales in the pelagic whaling ground of the North Pacific. Paper SC/SP74/Doc 23. Scientific Committee Special Meeting, La Jolla, 3-13 December 1974. 23pp.
Ohsumi, S. 1977. Further assessment of population of Bryde's whales in the North Pacific. Rep. int. Whal. Commn 27: 156-63.
Ohsumi, S. 1985. Estimation of the number of effectively marked Southern Hemisphere minke whales. Rep. int. Whal. Commn 35: 279-81.
Parker, R.A. 1963. On the estimation of population size, mortality and recruitment. Biometrics 19: 318-23.

Pollock, K.H. 1987. Mark and recapture techniques for estimating cetacean abundance. Paper SC/39/O6 presented to the IWC Scientific Committee, June 1987 (unpublished).
Seber, G.A.F. 1982. The Estimation of Animal Abundance and Related Parameters. 2nd ed. Griffin, London.
Shimadzu, Y. 1982. The unbalanced sex ratio in the minke whale catch and its effect on management. Rep. int. Whal. Commn 32: 905-9.
Shimadzu, Y. and Kasamatsu, F. 1984. Operating pattern of Antarctic minke whaling by the Japanese expedition in the 1982/83 season. Rep. int. Whal. Commn 34: 357-9.
Sigurjónsson, J. and Gunnlaugsson, T. 1985. Further mark-recapture analysis of fin whales caught off Iceland with a note on stock identity and movements of the east Greenland/Iceland population. Rep. int. Whal. Commn 35: 357-62.
Tillman, M.F. 1981. Estimated abundance of the western North Pacific stock of Bryde's whales. Rep. int. Whal. Commn 31: 385-8.
Tillman, M.F. and Breiwick, J.M. 1983. A note on mark-recapture estimates for North Pacific Bryde's whales. Rep. int. Whal. Commn 33: 443-5.
Tillman, M.F. and Grenfell, B. 1980. Estimates of abundance for the western North Pacific stock of Bryde's whales. Rep. int. Whal. Commn 30: 369-73.
Tillman, M.F. and Mizroch, S.A. 1982. Mark-recapture estimates of abundance for the western northern Pacific stock of Bryde's whales. Rep. int. Whal. Commn 32: 335-7.
Wada, S. 1984. Movements of marked minke whales in the Antarctic. Rep. int. Whal. Commn 34: 349-55.

## Note added in proof

Since this paper was written, the following corrections to the marking record have come to light:

## Mark no. 38579 in Table 4.

This is in fact Mark no. 38578, marked on 31/12/79 at $67^{\circ} 22^{\prime} \mathrm{S}, 40^{\circ} 12^{\prime} \mathrm{E}$ (verdict Ricochet). It was recovered on $19 / 1 / 86$ at $69^{\circ} 34^{\prime} \mathrm{S}, 0^{\circ} 41^{\prime} \mathrm{E}$ from an 8.9 m female.

Mark no. 39193 in Table 4.
This is in fact Mark no. 34193, marked on 12/1/79 at $67^{\circ} 54^{\prime} \mathrm{S}, 73^{\circ} 57^{\prime} \mathrm{E}$ (verdict Possible Hit). It was recovered on $4 / 4 / 87$ at $67^{\circ} 26^{\prime} \mathrm{S}, 75^{\circ} 05^{\prime} \mathrm{W}$ in the meat.

Mark no. 44236 in Table 4.
This is in fact Mark no. 44136, marked on 25/1/82 at $68^{\circ} 02^{\prime} \mathrm{S}, 20^{\circ} 02^{\prime} \mathrm{W}$ (verdict Possible Hit). It was recovered on $27 / 1 / 84$ at $68^{\circ} 20^{\prime} \mathrm{S}, 18^{\circ} 26^{\prime} \mathrm{E}$.

## Papers describing

 potential Management Procedures
# Simulation Studies of Two Whale Stock Management Procedures 

Justin G. Cooke<br>Center for Long Term Ecosystem Management, Moos, D-7809 Winden im Elztal 1, Federal Republic of Germany

Accepted April 1989


#### Abstract

Typical indices of abundance for whale stocks are not usually sufficiently precise to enable estimation of quantities such as maximum sustainable yield (MSY), replacement yield, or the maximum sustainable yield stock level (MSYL) within a reasonable time period However, even when these quantities are not known, it may be possible to design a combined assessment/management procedure which holds stocks near or above the MSY level while allowing sustained catches corresponding to a reasonable fraction of MSY to be taken, through continual adjustment of the catch based on the results of ongoing monitoring of the stock. Two potential management procedures are subject to preliminary evaluation using simulation trials. The initial results are promising, although they suggest that successful management of stocks which have already been severely depleted may be more difficult than the management of stocks which have only been lightly exploited prior to the implementation of the management regime.


## INTRODUCTION

The adoption of the New Management Procedure by the International Whaling Commission in 1974 was a major landmark in the development of management policies for marine resources. It supplemented the vague undertaking in the International Convention for the Regulation of Whaling to preserve whale stocks for future generations by a specific commitment to maintain whale stocks at population levels which yield close to their maximum sustainable yield (MSY). Stocks more than $10 \%$ below this level were to be protected, while a catch or exploitation rate corresponding to $90 \%$ of the MSY was to be taken from stocks at or above this level.

However, the implementation of the New Management Procedure was not generally successful in the attainment of this objective. The IWC Scientific Committee, which was entrusted with the task of estimating stock levels, MSY and MSY stock levels, was increasingly unable to provide the required estimates of the MSY and of the position of the stocks in relation to their MSY levels.

Many stocks were managed under an 'escape clause' in the New Management Procedure which permitted catches to remain at previous levels in the absence of positive evidence that the stock should be managed otherwise. 'Positive evidence' has been interpreted by the Scientific Committee to mean a statistically significant change in an index of abundance, such as Catch Per Unit Effort (CPUE), which was frequently the only available source of information on a stock (IWC, 1983a). De la Mare (1984) showed that typical levels of variability in CPUE series are such that a decline is not likely to be statistically significant until the stock is depleted to well below the protection level. Thus while such a procedure may serve to protect very depleted stocks, it will not lead to sustainable exploitation. To achieve sustainable exploitation, it is necessary to adjust catches before a stock declines into the protection category.
In the face of such problems, the IWC decided in 1982 to suspend those provisions of the NMP relating to catch limits and to replace them with interim catch limits of zero, effective from 1986, pending a reassessment of the situation by 1990 (IWC, 1983b).

## CAN SUSTAINABLE YIELDS BE ESTIMATED?

In the late 1970s and early 1980s the Scientific Committee estimated the sustainable yields of fin, sei and minke whales in the Southern Hemisphere from data on changes in the pregnancy rates and ages at first reproduction in the stocks (e.g. IWC, 1982), on the assumption that natural mortality rates do not change when a stock is depleted. However, reanalyses of the data and techniques used suggest that such apparent changes in pregnancy rates and age at maturity were artifacts (Mizroch and York, 1984; Cooke, 1985). Analyses by de la Mare (1986a) show that changes in true pregnancy rates are not observable from catch samples taken under realistic whaling conditions.

While samples of aged animals from catches provide some indication of the gross turnover in the population as a combination of natural mortality rates by age and trends in year class strength, they do not provide information on the relative contribution of these two factors. Thus they do not provide specific information on the net recruitment rate or sustainable yield. This point has been re-emphasised in recent discussions of the Scientific Committee on sampling programmes for age data: IWC (1989), de la Mare (1990), Cooke(1988), Nakamura, Ohnishi and Matsumiya (1989). Loosely speaking, age samples can provide an estimate of the sum of the net recruitment rate and the natural mortality rate, provided that assumptions can safely be made about the selectivity of the sampling process. Since the natural mortality rate cannot be less than zero, age samples may provide an upper bound on the possible value of the net recruitment rate.

When combined with independent data on the trend in population size, age data can be used to estimate the natural mortality rate and net recruitment rate separately. For the determination of sustainable yields, only the net recruitment rate is strictly relevant, and this can be estimated from the trend in population size alone: the age data do not contribute to the accuracy or precision of this estimate.

While data on population size and trends provide estimates of net recruitment rates and sustainable yields in principle, they may not be sufficiently precise to be directly usable. The surveys of minke whales in the Antarctic
conducted under the IDCR programme over the last eight years have yielded population estimates with a coefficient of variation of generally at least 0.2 , despite involving three or more vessels surveying only one-sixth part of the Antarctic over two months (IWC, 1988a).

If CV denotes the coefficient of variation of a population estimate from a single survey, then a programme of $n$ surveys conducted at regular intervals over a time span of $t$ years will yield a net recruitment rate estimate with standard error of approximately CV. $\left\{12 /\left(\mathrm{n} . \mathrm{t}^{2}\right)\right\}$.

At the most recent discussion of net recruitment rates of the Southern Hemisphere minke whales, opinions in the Scientific Committee ranged from under $2 \%$ to $4 \%$ (IWC, 1985). If the true net recruitment rate is $2 \%$, then to obtain an estimate of it accurate to within $\pm 20 \%$ at the $95 \%$ level would require 31 years of annual surveys with a CV of 0.2 . If surveys are only conducted once every six years, as has been the case for the Antarctic Areas under the IDCR programme, 76 years would be required.

The above calculations assume that the net recruitment rate is constant over the period. If the quantity of interest is the maximum sustainable yield (MSY), for which an estimate is required in order to implement the NMP, then the relationship between the net recruitment rate and population size needs to be determined, which may take even longer. In the case of an initially unexploited stock starting at carrying capacity, the net recruitment rate in the initial period of exploitation is relatively independent of the MSY exploitation rate.

Empirical determination of the MSY level requires measurement of the rate at which the net recruitment rate changes with changing population size. For this purpose it would be necessary to manipulate the stock size. A stable stock yields no information on the relationship between net recruitment rate and stock size. We could postulate an experiment in which a stock was reduced from $75 \%$ to $25 \%$ of its carrying capacity, and monitored continuously. If the true yield curve was a simple logistic with the MSY level at $50 \%$ of carrying capacity and an MSY rate of $2 \%$ of MSY stock size, then even from a 50 year experiment with annual estimation of stock size (with a CV of 0.2 ), the probability would be less than $50 \%$ that any significant relationship between net recruitment rate and stock size would be demonstrated. If the stock were depleted more quickly than this, the chance of detecting any density-dependence would be even less. The monitoring of the recovery of a depleted stock may be equally ineffective. If the stock were at $25 \%$ of its carrying capacity at the start of monitoring, then with an MSY rate of $2 \%$ it would reach about $70 \%$ of carrying capacity after 50 years, and the likelihood of detection of any density dependence would again be less than $50 \%$. Mere detection of density dependence would in any case be only a small step towards determination of the actual position of the MSY level.

If catch limits were based on direct estimates of replacement yields or MSY these would fluctuate greatly from year to year for several decades. Negative values would occur frequently, calling for a zero quota. This is confirmed by simulation studies by de la Mare (1986b), which indicate that it would take up to 100 years for catch limits to stabilise in a population managed on this basis. Whether the situation would actually stabilise after this time is somewhat doubtful, since over such long periods the population parameters may themselves change.

When only relative abundance data are available, the practice of the Scientific Committee has been to estimate
the absolute abundance by fitting a population model to the relative abundance series. This procedure will lead to even more unstable estimates of MSY, due to the correlation between estimates of absolute stock size and net recruitment rate. Again, this is confirmed by simulation studies by de la Mare (1986c,d).

It must be concluded that empirical determination of current or maximum sustainable yields is not possible in periods of less than 50 years or so. Determination of the MSY stock level may take much longer.

If measures to manage exploited whale stocks are to be limited to those based on the requirement to estimate the relevant quantities such as the MSY level and the MSY, then one is forced to conclude that there can be no scientific basis for management.

This does not necessarily exclude the possibility that whale stocks can be managed in such a way as to limit the risk of over-exploitation while at the same time obtaining a reasonable fraction of the maximum sustainable yield, even with very limited information on the true dynamics of the stocks. Such an approach to management will need to take account not only of the biological evidence for or against different assumptions about the level of sustainable yields, but also of the consequences of adopting incorrect values and of the time available to make corrections should they turn out to be wrong.

Two possible procedures for achieving this are outlined in this paper. Before going on to consider them, it may be helpful first to take note of some general principles of sustainable management.

The less frequently or accurately a stock is measured, the less the precision to which the stock can be held close to some optimal level. The yield that can be taken from, a stock without depleting it will depend therefore not only on the productivity of the stock itself, but also on the intensity of monitoring of the stock. There is a trade-off to be found between the resources devoted to monitoring a stock and the yield which can be taken. This is especially true if it is required to contain the risk of severely depleting the stock to some specified level: for a given level of risk, the catch that can be taken from a stock will be positively related to the intensity of monitoring because the characteristic time required to detect an adverse trend in a stock is inversely related to the intensity of monitoring.
There is also a trade-off to be sought between the average level of catch and the stability of the catch. A management procedure which aims to maximise the expected average yield subject to some stability constraint on the catches will achieve a lower average yield than one which maximises the average yield without any constraints. A management procedure aimed primarily at maximising the average catch, subject perhaps to some constraint on the risk of stock depletion, would tend to involve rapid adjustment of the catch level according to even weak evidence from the abundance data regarding the level of sustainable yield. A management procedure which placed a high premium on catch stability would involve only a sluggish reaction to the evidence from the data.

## A WEAKLY ADAPTIVE MANAGEMENT PROCEDURE

The first management procedure considered here shall be called the 'simple management procedure' to distinguish it from the slightly more adaptive and ambitious procedure in a subsequent section. The aim of the procedure is to hold
stocks near or above a notional MSY level, or allow them to recover to this level if they are already depleted; subject to this, reasonably stable catches corresponding to a reasonable fraction of the MSY are to be permitted. Further, those characteristics of the NMP which are considered to be workable are to be retained.

Leaving aside certain escape clauses, the main provision of the New Management Procedure was that catches should not exceed $90 \%$ of the MSY, and that stocks much below the MSY level should be protected to allow them to recover to productive levels as soon as possible. The IWC Scientific Committee recognised at an early stage the impossibility of estimating the MSY population directly, and instead assumed a notional level of first $50 \%$, subsequently $60 \%$, of carrying capacity for the MSY level. According to the standard Pella-Tomlinson yield curve used by the Scientific Committee to represent the notional relationship between population size and sustainable yield, a whale stock exploited with an MSY level at $60 \%$ of carrying capacity ( K ) from which constant catches corresponding to $90 \%$ of the MSY are taken, would theoretically stabilise at approximately $75 \%$ (strictly, $73.8 \%$ ) of carrying capacity. In the absence of evidence to the contrary, carrying capacity has been interpreted to mean the notional population level prior to the start of human exploitation.

The procedure proposed here is also to take $90 \%$ of the MSY from a stock which is above the target level of $75 \%$ of carrying capacity ( 0.75 K ), but to reduce this catch level for populations below this level to zero for populations at 0.5 K of carrying capacity according to the control law shown in Fig. 1. Since the true value of the MSY is not known, it is defined to be a fixed percentage of the MSY population size. In the results shown below, the value of $3.5 \%$ was chosen to give a reasonable compromise between the risk of stock depletion and the level of catch taken.

Because of the variability in the population estimates used to assess the status of the stock, a stock near the protection level may be quite likely to be protected and de-protected several times in succession. To reduce the likelihood of this happening, stocks which have been protected are not re-opened for exploitation until they are estimated have recovered to at least 0.6 K .

To implement the procedure requires an estimate of current population size and of the ratio of the current population size to the notional unexploited level or carrying capacity ( K ). The value used for K need not necessarily represent an actual stock level that pertained at any particular time in the past.


Fig. 1. Control law for the simple management procedure: catch limit as a fraction of MSY as a function of stock size as a fraction of K.

The current population level is estimated directly from surveys of the population size. If surveys are conducted regularly, then the variance of the stock estimate can be reduced by averaging the estimates obtained from surveys over several years. The longer the period over which estimates are averaged, the less the variability of the estimates obtained, but the greater the possible bias arising from changes in the population level over the period of averaging. In this procedure it is proposed to leave the choice of the numbers of years averaged free, but to discount estimates from surveys in previous years by at least $2 \%$ of the estimate each year, or the amount of catch taken, whichever is the greater. The lower $95 \%$ confidence limit of the resulting averaged estimate is then used as the current population estimate, in order both to provide some protection against over-exploitation that could arise from the uncertainty in the estimate, and to attach a positive value (in terms of catch permitted) to the precision of the estimate. It is assumed in the simulations that follow that the averaging period is chosen each year so as to maximize the resulting catch quota. The use of the lower $95 \%$ confidence limit tends to make estimates based on a longer averaging time larger due to their lower variance, while the discounting procedure tends to reduce them. These two opposing influences will result in a maximum population estimate at some intermediate length of averaging period.

Population estimates not included within the averaging period for the estimation of current population size are discarded.

Estimates of population size from surveys are normally accompanied by nominal estimates of their variance, although such nominal variances may be negatively biassed with respect to the true variance because they do not take account of all possible sources of variability in the estimates. The empirical variance of the sample of population estimates used to obtain the averaged estimate also provides some information on the variance, although such empirical variance estimates are themselves highly variable and will in a proportion of cases give a misleading impression of precision when successive population estimates are by chance very similar. In the simulations that follow, the lower $95 \%$ confidence limits have been calculated either from the nominal variances (assumed known) or from the empirical variance of the sample of estimates over the averaging period, according to which is the greater.

The unexploited population level is estimated by 'shooting' a simplified Pella-Tomlinson population model through the current population estimate. The dynamic equation of the model is:

$$
\begin{equation*}
N_{t+1}=N_{t}\left(1+r\left(1-\left(N_{t} / K\right)^{z}\right)\right)-C_{t} \tag{1}
\end{equation*}
$$

where $N_{t}$ is the population size in numbers in year $t, r$ is the intrinsic growth rate, $K=N_{o}$ is the unexploited population level, and $z$ is the density-dependent exponent. The conventional choice of the MSY level at 0.6 K corresponds to a value of $z$ of 2.39. The MSY is the following fraction of MSY stock level: $\mathrm{rz} /(1+\mathrm{z}) . \mathrm{C}_{\mathrm{t}}$ is the catch n year t .

The time lag due to the non-zero age at maturity or recruitment, and the natural mortality rate, have been omitted from the population model, partly to simplify the calculations and partly to emphasise the relatively minor importance of these parameters.

Of the three parameters in the model ( $\mathrm{r}, \mathrm{K}$, and z ) r and z are assigned their assumed values. Only K is estimated by fitting the model through the current population estimate.


Fig. 2. Long term equilibrium stock size (as a fraction of K ) from deterministic simulation of the simple management procedure, as a function of the true MSY rate, for an assumed MSY rate of $3.5 \%$.

Population estimates from before the averaging period are ignored in the fit. The procedure is thus at best only weakly adaptive. If the population is declining, the current population estimate will decline and hence the quota. The estimated ratio of current to initial population size will also decline, but to a lesser extent, since the estimate of $K$ will also decline somewhat. Hence, once the population is estimated to be below the target level, the exploitation rate will decline and the population will eventually stabilise, even if the true MSY rate is below the assumed value. Fig. 2 shows the proportion of K at which the population would theoretically stabilise in a deterministic simulation of the procedure, as a function of the true MSY rate for an assumed MSY rate of $3.5 \%$. It would appear that if the true MSY rate is much below the assumed value, the population could be severely depleted. However, the rate of depletion slows down as the stock is depleted. Thus while the long term stable population level under the procedure may be far from the MSY level, the procedure may be acceptable for the first 100 years of exploitation, after which the evidence from the data could be used to adjust the assumed MSY rate.

To reduce undesirable fluctuations in catch levels from one year to the next, the catch in each year is constrained to lie within $20 \%$ of the previous year's catch, unless the latter is less than 20 animals. This rule is recommended by IWC (1989).

Given the evident impracticability of testing management procedures on real stocks (cf. the time scales required to obtain informative results, the risks to the stocks entailed by the tests, the large number of test stocks required in order to assess the probability of outcomes which are possible but not inevitable), it is clear that simulation studies represent the primary means available for evaluation of procedures.

The IWC Comprehensive Assessment Workshop on Management (IWC, 1988b) specified certain initial simulation trials to which any proposed whale stock management procedure should first be subjected. This specification was slightly amended at the second workshop (IWC, 1989).

Computer simulations of whale stock management procedures enable evaluation of the performance under conditions that can be envisaged in advance. Satisfactory performance in such simulation trials does not necessarily imply that the procedure would perform well in reality, because factors may operate which were not foreseen in
the design of the simulation trials. The converse, however, is more convincing: if a management procedure does not perform adequately in the relative simple and idealized conditions of a simulation trials, it is hardly likely to fare better in practice. Simulation trials at least have the potential to 'weed out' some of the less desirable characteristics of management procedures.

A full specification of the simulation trials is to be found in IWC (1989). Briefly, the trials are to be conducted for all combinations of the following situations, yielding 8 cases in all: current stock size estimates are assumed to become available every fifth year with a coefficient of variation (CV) of (i) 0.2 or (ii) 0.4 ; the true population is assumed to follow a Pella-Tomlinson model with MSY level at 0.6 K with a true annual MSY rate of (i) $1 \%$ (ii) $4 \%$; the true population is initially unexploited at the start of management or already depleted to 0.3 K by 30 years of constant catch. The management procedure is to be simulated for 100 years for each case. In the case of the initially depleted population, the previous catch history is known, but no population estimates are available for the pre-management period. This is typical of the situation of most whale stocks that have been exploited to date. The age at maturity in the true population is 7 years and the natural mortality rate is $0.05 \mathrm{yr}^{-1}$.

The unexploited population level in the simulation trials is 10,000 , hence the MSY is 60 and 240 for the $1 \%$ and $4 \%$ cases respectively. In accordance with the terminology used by de la Mare (1989) and IWC (1989), the case of the stock without a previous history of exploitation is called the development case, while the case of the stock depleted before the start of management is called the rehabilitation case.

The simulation trials are designed to reflect the true management situation in its essential structure: although the true population is simulated, the only information on the true population that the management procedure may make use of are the 5 -yearly stock estimates which are subject to the above degrees of random error.

In this paper, only the results of trials assuming 5-yearly population estimates with a CV of 0.2 are presented. It is intended to present the results for a CV of 0.4 in a later paper after further tuning of the procedure. This leaves four cases, which shall be labelled D1, D4, R1 and R4, denoting the development and rehabilitation cases for true MSY rates of $1 \%$ and $4 \%$ respectively.

Table 1 gives the following statistics from 100 independent random simulation trials of 100 -year runs of the procedure. These include the statistics identified by IWC (1990) as required for a preliminary evaluation of potential management procedures.

## Table 1

Summary statistics from 100 simulation trials of the simple management procedure

| Case code | Minimum pop. |  | Lowest min. pop. | Final pop. |  | Average catch |  | Average yield |  | R.m.s. quota change | Av. <br> SD of <br> catch |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Mean | SD |  | Mean | SD | Mean | SD | Mean | SD |  |  |
| D1. 2 | 5,443 | 216 | 4,836 | 5,455 | 214 | 108.8 | 3.1 | 62.8 | 1.3 | 8.0 | 27.1 |
| R1. 2 | 3,000 | 0 | 3,000 | 6,113 | 408 | 13.6 | 5.5 | 45.0 | 1.4 | 12.2 | 27.8 |
| D4.2 | 8,470 | 112 | 8,142 | 8,691 | 159 | 137.6 | 5.3 | 124.3 | 4.8 | 9.4 | 22.7 |
| R4.2 | 3,000 | 0 | 3,000 | 9,200 | 350 | 58.1 | 10.5 | 120.7 | 9.1 | 19.6 | 58.8 |

(i) the mean and standard deviation of the lowest population size in 100 years;
(ii) the lowest population level reached in any trial;
(iii) the mean and standard deviation of the population size after 100 years of management;
(iv) the mean and standard deviation of the mean catch over 100 years;
(v) the mean and standard deviation of the mean yield over 100 years;
(vi) the root-mean-square change in catch from one year to the next;
(vii) the average of the standard deviation of the catch within each trial;
Statistics (i) and (ii) provide some indication of the risk of depletion of the stock and the ability of the procedure to prevent depletion of the stock to far below the MSY level. Perhaps the most relevant measure of risk is the lowest population achieved in any trial, but it should be recognised that this statistic is to a large extent a random result of the particular set of trials conducted, and a further set of 100 similar trials could lead to a substantially different value.

In the rehabilitation cases the lowest population size statistics are not especially informative because in no case was the stock reduced below its initial, depleted level. The population size after 100 years gives an indication of the ability of the procedure to rehabilitate the stock to productive levels.

The mean catch taken gives some indication of the ability of the procedure to utilise the true productivity of the stock over the period. However, in the development case the initial depletion of the stock 'capital' makes a substantial contribution to the mean catch. Since a large catch can easily be achieved by depleting a stock, a possibly more appropriate measure of the true ability of the procedure to realise the productivity of the stock is the total yield taken. Yield is defined here to be the catch taken adjusted by the change in population size. The mean yield over the 100 -year period is given by:

$$
\text { mean yield }=\text { mean catch }+(\text { change in stock size }) / 100
$$

The mean yield is less than the mean catch in the development case, because the change in stock size over the period is negative. The mean yield is greater than the mean catch in the rehabilitation case, because the change in stock size is positive. Because of the time lag arising from the age at maturity ( 7 years) in the model used to generate the true population dynamics, it is possible for the yield obtained to exceed the MSY slightly over a finite period, although to a good approximation the MSY can be regarded as an upper bound on the feasible mean yield.

The root-mean-square change in the catch level from year to year was originally chosen as a statistic to reflect the stability of catch limits implied by the procedure (IWC, 1988b). However, it it is not an entirely satisfactory measure of stability because it assigns an equal penalty to small fluctuations around a stable level as it does to large changes over periods of several years provided that the latter are made in small annual steps. Therefore, a further statistic, the standard deviation of catch over the 100-year period was prescribed as an additional statistic on catch stability to be selected (IWC, 1989).

Figs 3-6 show the mean and range of (a) population sizes and (b) catch by year from the 100 trials, for each of the four cases. Some features of these results are discussed overleaf:

(a)

(b)

Fig. 3. Mean and range of a) stock size and b) catch from 100 simulation trials of the simple management procedure. Development case with MSY $=1 \%$.

(a)

(b)

Fig. 4. Mean and range of a) stock size and b) catch from 100 simulation trials of the simple management procedure. Rehabilitation case with MSY $=1 \%$.


Fig. 5. Mean and range of a) stock size and b) catch from 100 simulation trials of the simple management procedure. Development case with MSY $=4 \%$.


Fig. 6. Mean and range of a) stock size and b) catch from 100 simulation trials of the simple management procedure. Rehabilitation case with MSY $=4 \%$.

## Development case, MSY = 1\%

The population is reduced steadily to approximately the MSY level after 100 years, with relatively little variation across the 100 scenarios (Fig 3a). Inevitably, the catch varies more between scenarios than the stock (Fig. 3b). The mean catch is still slightly above the MSY after 100 years, hence a transition to a more responsive management procedure would be necessary to prevent further declines in the stock after this time. The lowest population reached in any trial is 0.48 K which suggests that the risk of severe depletion may be acceptably small.

## Rehabilitation case, MSY=1\%

Even without catches, the stock does not recover to the formal target level of 0.75 K within 100 years, and only reaches the MSY level ( 0.6 K ) after 77 years. The mean yield would therefore be maximised by taking no catch in the first 77 years. The results shown in Figs $4 a-b$ show that on average the population is allowed to recover to just over the MSY level, but in some simulations no catch is taken at all, allowing the stock to reach 0.7 K . The stock is rehabilitated to at least 0.5 K in all trials. In no case is the stock depleted below the level pertaining at the start of management.

## Development case, MSY=4\%

The stock is in no case depleted to below 0.8 K at any time. Average catches soon stabilise at about $60 \%$ of MSY, which is two-thirds of the notional target catch level of $90 \%$ of MSY. The stock is under-utilised relative to the target catch, but possibly not to an unacceptable degree. In no cases do catches exceed MSY.

## Rehabilitation case, MSY=4\%

The procedure is successful in rehabilitating the stock to above the target level, but in most simulations the full potential yield is not realised. Average catches stabilise at about $40 \%$ of MSY after the stock has recovered to above the target level. The performance in this case is in some respects the least satisfactory of the four cases.

## AN ADAPTIVE MANAGEMENT PROCEDURE

Given the failure of the above simple procedure to realise the full potential of the stock in some cases, in particular in the case of an initially depleted stock whose true potential productivity level is high, it may be worth considering slightly more adaptive procedures which make more use of the abundance data to estimate the productivity of the stock. The considerations earlier in this paper suggest, however, that the information content of the data is relatively low and hence that the scope for improvement over the simple procedure may be rather limited.
The procedure prosed here is to estimate both the K and $r$ parameters in the population dynamic model (1), while continuing to use the assumed value of $z$. Furthermore, in each year all the abundance estimates will be used to estimate the parameters rather than merely the most recent ones.
An immediate problem that arises when attempting to estimate $r$ by fitting the model to abundance data is that such estimates are numerically badly behaved. For example, for the initially unexploited stock, it is not unlikely that in the first few years the data will exhibit a (non-significant) increasing trend in the population even when the true population has a small decreasing trend.

Since the model does not allow an unexploited population to increase, the best-fitting estimate of $r$ from such data would be infinite, which corresponds to a population that is so resilient that no catches can deplete it. To overcome this bad behaviour, the model can be re-parameterised as follows:

$$
\begin{equation*}
N_{t+1}=N_{t}\left(1+r_{o}(\alpha /(1-\alpha))\left(1-\left(N_{t} / K\right)^{z}\right)-C_{t}\right. \tag{2}
\end{equation*}
$$

where $\alpha=r /\left(r+r_{o}\right)(0 \leq \alpha<1)$
$\alpha$ can be called the resistance of the stock. $\alpha=0$ implies $r=$ 0 and that no catch from the stock is indefinitely sustainable. $\alpha=1$ implies that the stock is not affected by any level of catch. $r_{o}$ is a reference value of $r$ corresponding to $\alpha=0.5$. In the simulations that follow a value of 0.02 was used for $r_{0}$.
$\alpha$ and K are jointly estimated from a maxi-mum-likelihood fit of the population model (2) to the series of population estimates using the known catches. The maximum likelihood estimate of $\alpha$ is not used directly Instead, the mid-point of the $95 \%$ confidence limits for $\alpha$ is used. Having selected this value of $\alpha, \mathrm{K}$ is re-estimated from a maximum-likelihood fit to the data keeping $\alpha$ fixed at this value. The resulting fitted population trajectory also provides an estimate for the current population size, $\mathrm{N}_{\mathrm{t}}$.

## Confidence interval calculation

Exact confidence limits for $\alpha$ are not readily calculated, nor is it especially important to do so. When there are three or more data points, then by analogy with linear theory, the variance of the population estimates can be estimated empirically from the residual deviations from the fitted model: the corresponding critical values of the F variance ratio statistic can be used to obtain nominal confidence limits. De la Mare (1989) shows that that confidence intervals obtained using the F-ratio statistic for the estimated population depletion from a similar model contain the true population depletion value with a probability close to the nominal $95 \%$ value. However, the widths of confidence intervals from the $F$ ratio method are highly stochastic when there are relatively few data points. Population estimates are usually accompanied with nominal estimates of their variance. Even though these may be negatively biassed through failure to incorporate all sources of variability, they nevertheless provide some relevant information. In particular if the empirical variances of the population estimates as estimated from the size of the residual deviations from the model fit are lower than the calculated variance estimates, then the manager would be aware that use of the empirical variance estimates would provide a false impression of accuracy. The relative values of the nominal population estimate variances - or some related statistic such as survey effort or numbers of animals seen - are in any case required to determine the relative prior weights to be given to each data point even when the F-ratio method is used.

In this procedure, a combined approach has been adopted. The empirical variance estimates from the fit of the model are used where these are greater than the nominal variances of the population estimates, otherwise the nominal variances are used. When the nominal values are used, confidence limits are obtained from the corresponding percentage points of the chi-squared distribution. This selection procedure will tend to bias the estimate of variances upwards, resulting in conservative confidence intervals. The nominal variance is the only one available when there are less than three data points.


Fig. 7. Control law for the adaptive management procedure: fishing mortality limit as a fraction of the MSY fishing mortality, as a function of stock size as a fraction of $K$

From the results presented below, it is apparent that the transition to the F-ratio method which usually occurs in year 15 of the simulations with the arrival of the third data point, has led in the majority of cases to an exaggerated assessment of uncertainty and hence unnecessarily conservative catch limits in years $15-20$. In a later refinement of this method it is intended to retain the selection of the higher of the two variance estimates, but to use the chi-squared method of estimating confidence limits in each case as if the variances were known. This would be expected to remove the anomaly that arises when there are exactly three estimates. An alternative would be use the F-ratio method only when there are at least, say, five population estimates.

As the number of data points increases, the confidence limits for $\alpha$ tend towards symmetry about the estimate of $\alpha$, and so the procedure approaches a certainty-equivalent procedure. With fewer data, the confidence limits are asymmetrical about the estimate. Initially, the confidence interval for $\alpha$ is the range $(0,1)$, so that the value of $\alpha$ used is 0.5 , which corresponds to an $r$ value equal to $r_{0}$. Only when there are sufficient data for the confidence limits for $\alpha$ to exclude one or both ends of the $(0,1)$ range is a different value of $r$ used.

As in the case of the simple management procedure, the protection level is set at $50 \%$ of the estimated K value, but exploitation of stocks which have already been protected is not re-opened until the estimated stock size exceeds 0.6 K . Above these levels, the fishing mortality is set to the fraction of the MSY fishing mortality implied by the control law shown in Fig. 7. The maximum value of the fishing mortality rate of 0.7 times the MSY fishing mortality rate is the value that would theoretically stablise the stock at approximately 0.75 K , at which level the catch would be $90 \%$ of MSY. This accords with the objectives of the simple management procedure and the NMP.

As a further measure to take account of the uncertainty in the estimate of the current stock size in relation to $\mathbf{K}$, a multiplicative adjustment factor based on the lower confidence limit of the estimated stock depletion (ratio of current stock size to $K$ ) is applied to the catch. If $D_{L}$ is the lower $95 \%$ confidence limit of the estimate of depletion, the adjustment factor is:

$$
\begin{gathered}
0\left(\mathrm{D}_{\mathrm{L}} \leq 0.25\right) \\
4(\mathrm{D}-0.25)\left(0.25<\mathrm{D}_{\mathrm{L}} \leq 0.5\right) \\
1\left(\mathrm{D}_{\mathrm{L}}>0.5\right)
\end{gathered}
$$

Table 2
Summary statistics from 100 simulation trials of the adaptive management procedure

| Case code | Minimum pop. |  | Lowest min. pop. | Final pop. |  | Average catch |  | Average yield |  | R.m.s. quota change | Av SD of catch |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Mean | SD |  | Mean | SD | Mean | SD | Mean | SD |  |  |
| D1.2 | 5,935 | , 022 | 3, 712 | 6,436 | 1,176 | 92.4 | 16.2 | 56.4 | 4.5 | 10.2 | 49.5 |
| R1.2 | 3, 000 | 0 | 3, 000 | 5,488 | 490 | 21.7 | 6.3 | 46.8 | 1.4 | 6.1 | 26.9 |
| D4.2 | 6,885 | 587 | 5,646 | 7,310 | 762 | 169.1 | 20.0 | 142.0 | 14.9 | 16.9 | 74.7 |
| R4.2 | 3,000 | 0 | 3,000 | 8,310 | 567 | 106.5 | 33.8 | 160.1 | 28.4 | 13.8 | 75.6 |


(a)

(b)

Fig. 8. Mean and range of a) stock size and b) catch from 100 simulation trials of the adaptive management procedure. Development case with MSY $=1 \%$.

Thus there is no adjustment when the lower confidence limit of the estimate of stock depletion is above the protection level, and no catch is allowed if the lower confidence limit is below half the protection level.

Finally, the same catch fluctuation control rule is applied as was used with the simple management procedure.

Table 2 and Figs 8-11 give the corresponding results to those given for the simple management procedure in Table 1 and Figs 3-6.

## MSY $=\mathbf{1 \%}$ : development and rehabilitation cases (Figs 8-9)

Although the mean stock trajectories are satisfactory, stock sizes as low as 0.37 K occur in trials of the development case. In the rehabilitation case with MSY = $1 \%$, some trials failed to rehabilitate the stock to above 0.4 K , although no stock was depleted to below the initial level of 0.3 K . The spread of stock trajectories across trials is much greater than for the simple management procedure. For neither of the $1 \%$ MSY cases does the adaptive procedure offer any clear advantage over the simple procedure.

(a)

(b)

Fig. 9. Mean and range of a) stock size and b) catch from 100 simulation trials of the adaptive management procedure. Rehabilitation case with MSY $=1 \%$.

(a)

(b)

Fig. 10. Mean and range of a) stock size and b) catch from 100 simulation trials of the adaptive management procedure. Development case with MSY $=4 \%$.

MSY $=4 \%$ : development and rehabilitation cases (Figs 10-11)
In terms of the mean catches and population size, the problem of under-utilisation of the stock with a 4\% MSY rate that afflicted the simple procedure has been largely


Fig. 11. Mean and range of a) stock size and b) catch from 100 simulation trials of the adaptive management procedure. Rehabilitation case with MSY $=4 \%$.
solved. For the 4\% MSY case, mean catches stablise at around $85-90 \%$ of MSY in the development case and $65-70 \%$ of MSY in the rehabilitation case. However, this is at the price of considerably increased variability. Despite the improvement in the mean catch in the development case, the minimum catches are slightly lower than for the simple management procedure. As with the simple management procedure, exploitation in the rehabilitation case is not re-opened in some trials until the stock has recovered to close to K , despite the reasonably good mean catch level. All trials left the stock above the MSY level after 100 years in both the development and rehabilitation cases.

## DISCUSSION

The adaptive management procedure clearly requires some further development. In particular the anomaly at year 15 already referred to should be removed. Another anomaly is that in the development case the initial catch is a fixed proportion of the initial population estimate regardless of the variance of the latter. The tendency of the catch to decrease in the first few years of the development case could probably be solved by replacing the control law based on fishing mortality as a fraction of the MSY fishing mortality by one based on the catch as a fraction of the MSY catch. This would also enhance comparability with the simple management procedure.

It should be emphasised that the simulation results presented in this paper only address certain aspects of the management procedure, namely the procedure for setting catch limits as a function of the observations of stock size. Issues such as the designation of management stocks or areas have not been addressed at all. Some of these other issues are to be addressed in the second set of simulation trials specified by IWC (1989).

In some respects the management procedures considered here resemble closely the New Management Procedure adopted by the IWC in 1974. Both are based on notions of quantities such as the MSY and MSY population levels, and both use control laws of a similar nature. However, the approaches outlined here differ in one fundamental aspect from the NMP. The NMP was phrased in terms of the MSY and MSY level as objective biological characteristics of stocks that had to be estimated. Since these could not be estimated, the NMP was largely unsuccessful in achieving sustainable exploitation. The approaches developed here treat the MSY and MSY level as purely notional quantities used to determine catch limits. The catch limits are derived in a direct mechanical fashion from the data obtained: there is no scope for argument about the biological correctness or otherwise of the parameter values used, nor about whether the catch limits set by the procedure are justified by the data available. The procedures even set catch limits in cases where the only information available is an estimate of absolute stock size, without any information on the sustainable or replacement yield. The main criterion for choosing the parameter values in the models used is not the relative biological plausibility of the values chosen, but the consequences of adopting them even if they turn out to be wrong.

The catch limits set in individual years by the procedure of the kinds presented here are not claimed to be based on specific scientific evidence, and are justified only in the context of the performance of the given procedure as a whole. Even if the procedure performs satisfactorily on average, there will be years when catch limits are 'unnecessarily' raised or lowered. The procedures involve the deliberate abandonment of the conventional requirement that the need for each specified management action be scientifically demonstrated before any such action can be taken. It may be that this is a necessary price to pay if sustainable exploitation of whale stocks is to become a feasible objective.

The procedures examined here share this characteristic with those examined by de la Mare (1989), Butterworth and Punt (1989) and Sakuramoto and Tanaka (1989).

## ACKNOWLEDGEMENTS

W.K. de la Mare, Marine and Ecological Research, Beaumaris, Victoria, Australia, and S. J. Holt, Podere 'Il Falco' (Loc. Acquaioli), Cittá della Pieve, I-06062 Perugia, Italy, kindly reviewed the manuscript.

## REFERENCES

Cooke, J.G. 1985. Have Southern Hemisphere minke whale ages at maturity declined? Rep. int. Whal. Commn 35:335-40.
Cooke, J.G. 1988. Comments on the proposed Scientific Permit for the take of minke whales in Antarctic Area V. Paper SC/3 circulated by IWC, November 1988 (unpublished).
de la Mare, W.K. 1984. On the power of catch per until effort series to detect declines in whale stocks. Rep. int. Whal. Commn. 34:655-62.
de la Mare, W.K. 1986a. Inferring net recruitment rates from changes in demographic parameters: a sensitivity analysis. Paper SC/38/O 2 presented to the IWC Scientific Committee, May 1986 (unpublished). 12pp.
de la Mare, W.K. 1986b. Simulation studies on management procedures. Rep. int. Whal. Commn 36:429-50.
de la Mare, W.K. 1986c. Further consideration of the statistical properties of catch and effort data, with particular reference to fitting population models to indices of relative abundance. Rep. int. Whal. Commn 36:419-24.
de la Mare, W.K. 1986d. Fitting population models to time series of abundance data. Rep. int. Whal. C'ommn 36:399-418.
de la Mare, W.K. 1989. Simulation studies of elements of a revised whale management procedure. (Published in this volume).
de la Mare, W.K. 1990. A further note on the simultaneous estimation of natural mortality rate and population trend from catch-at-age data. Paper SC/41/O 1, to be published in Rep. int. Whal. Commn. 40.

Intcrnational Whaling Commission. 1982. Report of the Special Meeting on Southern Hemisphere minke whales. Rep. int. Whal. Commn 32:697-746.
International Whaling Commission. 1983a. Report of the Scientific Committee, Agenda Item 8.3.1. Rep. int. Whal. Commn 33:48.
International Whaling Commission. 1983b. Chairman's Report of the 34th Annual Meeting. Rep. int. Whal. Commn 33:20-42.
International Whaling Commission. 1985. Report of the Scientific Committee, Agenda Item 10.2.7. Rep. int. Whal. Commn 35:42-3.
International Whaling Commission. 1988a. Report of the sub-committee on Southern Hemisphere minke whales. Appendix 7. Revised estimates of takeable minke abundance from the 1978/791983/84 Antarctic IDCR sightings cruises. Rep. int. Whal. Commn 38:86-87.

International Whaling Commission. 1988b. Report of the Comprehensive Assessment workshop on management, Reykjavík, March 1987. Rep. int. Whal. Commn. 88: 163-70 [and reprinted in this volume].
International Whaling Commission. 1989. Report of the Scientific Committee. Annex J5. Working group to develop a protocol for the evaluation of procedures to estimate population dynamics parameters from age data. Rep. int. Whal. Commn 39: 136-8
International Whaling Commission. 1989. Report of the Comprehensive Assessment workshop on management procedures, Lowestoft, UK, 6-10 February 1989. (Published in this volume).
Mizroch, S.A. and York, A.E. 1984. Have pregnancy rates of Southern Hemisphere fin whales, Balaenoptera physalus, increased? Rep. int. Whal. Commn (special issue 6):401-10.
Nakamura, T., Ohnishi, S. and Matsumiya, Y. 1989. A Bayesian cohort model for catch-at-age data obtained from research takes of whales. Rep. int. Whal. Commn. 39: 375-82.
Punt, A.E. and Butterworth, D.S. 1989. Results of first stage screening trials for a proposed whale stock management procedure. (Published in this volume).
Sakuramoto, K., and Tanaka, S. 1989. A simulation study on management of whale stocks considering feedback systems. (Published in this volume).

# Further Simulation Studies on Management Procedures 

William K. de la Mare<br>Department of Zoology, Monash University, Wellington Road, Clayton, Victoria, Australia.

Accepted September 1988


#### Abstract

In this paper, further simulation studies on a revised management procedure are reported. The studies here extend earlier simulation studies on management procedures, by examining the performance of the revised procedure (RMP) described in de la Mare (1986a and $b$ ), with biased estimates of absolute abundance. This is found to lead to a failure in achieving management goals. Treating the absolute abundance data as an index of relative abundance overcomes this problem without leading to the problem which arises in the catch-effort case of failing to realise the potential of the stocks. A preliminary investigation is made of a modification to the RMP which adjusts the catch limits according to the uncertainty in parameter estimates.


## INTRODUCTION

A series of simulation studies on the properties of management procedures was reported in de la Mare (1986a). In that study, the New Management Procedure (NMP) and a revised management procedure (RMP) were tested when the estimates of parameters required to implement them were estimated by means of fitting population models to times series of either relative or absolute abundance data. The RMP was based on the incorporation of explicit feedback, such that if the population were depleted to below an arbitrary target level, then catches were set at less than the replacement yield (RY), and more than the RY if the population is above the target level. The RMP procedure was found to give a better performance than the NMP in terms of maintaining an exploited population near the target level, and in particular, above the protection level.

It was found that fitting models to catch-effort data led to a tendency to drift to low levels of exploitation, but this problem did not arise when management was based on models fitted directly to unbiased estimates of absolute abundance. The postulated explanation for this was the lack of separation between the catches and information about the status of the stocks, when using catch-effort data. In this paper, further confirmation of this explanation is obtained. The studies reported here are restricted to the case where there is separation of information and catches.

## THE SIMULATION MODEL

The performance of the RMP is examined by means of simulation trials of the complete management process. The simulation model used is that described in de la Mare (1986a). A two component population model is used to generate the time series of absolute abundance data. Assessments are made by fitting the same form of population model to the absolute abundance data, although the fitted model usually has erroneous parameter values. The fitted population model acts as a 'filter' which separates the signals required by the management procedure from the 'noisy' data. The population model is given by the following expressions:

$$
\begin{equation*}
P_{t+1}=\left(P_{t}-C_{t}\right) S+R_{t} \tag{1}
\end{equation*}
$$

where $P_{t}$ is the exploitable population size in year $t, S$ is the natural survivorship, and $R_{t}$ is the recruitment given by:

$$
\begin{equation*}
\mathrm{R}_{\mathrm{t}}=\mathrm{P}_{\mathrm{t}-\mathrm{m}}(1-\mathrm{S})\left\{1+\mathrm{A}\left[1-\left(\mathrm{P}_{\mathrm{t}-\mathrm{m}} / \mathrm{P}\right)^{\mathrm{z}}\right]\right\} \tag{2}
\end{equation*}
$$

Table 1
Parameter values used in Equation 2

| Population parameter set 1 | Population parameter set 2 |
| ---: | :---: |
| $\mathrm{m}=7$ | $\mathrm{~m}=7$ |
| $\mathrm{~A}=0.1984$ | $\mathrm{~A}=0.7824$ |
| $\mathrm{~S}=0.9324$ | $\mathrm{~S}=0.9324$ |
| $\mathrm{z}=4.04$ | $\mathrm{z}=1.39$ |
| MSYL $=0.7$ | MSYL $=0.6$ |
| MSY $=1.2 \%$ (per capita) | MSY $=4.0 \%$ (per capita) |

where $m$ is the age at first parturition, $A$ is the range of density dependent response, $P$ is the size of the population prior to exploitation and $z$ is the density dependent exponent. Equation (2) incorporates a balance equation. The two sets of parameter values used, chosen to represent a plausible situation, are given in Table 1.

The emphasis of the trials is on the properties of the RMP when the fitted model overestimates per capita MSY. Thus, in most trials, the model for the true population uses parameter set (1), and the fitted population model uses parameter set (2). The initial estimates for the fitted population are virtually the 'standard' values used in assessment over most of the life of the NMP. The 'true' population parameters give a per capita MSY towards the low end of the range considered likely by the Scientific Committee of the IWC (IWC, 1983).

The initial catch is calculated as $90 \%$ of the estimate of MSY from the fitted model using an initial estimate of absolute abundance. After the initial period, catch limits are set from the parameters of the fitted model using the management control law. Thus, the catches and apparent population trajectories are not independent. The parameters chosen as measures of performance are the distributions of true population size after given periods of exploitation, the distribution of catches taken and the distributions of estimates of population parameters, particularly depletion ( D , the ratio of final to initial population size). Catch limits are never set at more than twice their average value prior to the time of assessment. Thus, assessments which suggest large increases in catch limits are ignored. The concept of 'certainty equivalence' is applied, that is, the estimates of parameters are acted on at each stage as if they were without error (certainty equivalence is reviewed in Goodwin and $\operatorname{Sin}$, 1984).

[^11]The management control law used is described fully in de la Mare (1986a). In essence, catch limits are set at a proportion of the estimate of RY, the proportion depending on the difference between the estimated depletion of the stock and a target level. If the stock is above the target level, catches are greater then the RY, and conversely, if the stock is below target then catches are less than RY. The exact form of the control law can be expressed as follows:

$$
\begin{align*}
\mathrm{G} & =(\mathrm{D}-\mathrm{Q}) /(\mathrm{T}-\mathrm{Q}), & \mathrm{D} & >\mathrm{Q} \\
& =0, & & \mathrm{D} \leq \mathrm{Q} \tag{3}
\end{align*}
$$

where D is the estimated depletion, Q is the protection level and T is the target level. In these trials, T is $75 \%$ of initial population size, and $Q$ is set at $55 \%$. These parameters are close to the conventional levels of depletion which would arise from the application of the NMP with MSY occurring at $60 \%$ of initial population size. The catch limit is given by:

$$
\begin{equation*}
\mathrm{C}=\widehat{\mathrm{RY}} \cdot \mathrm{G} \tag{4}
\end{equation*}
$$

Simulation trials are carried out, using a fitted population model to estimate the depletion and RY. The fitting criteria used are the same as for the trials in de la Mare (1986a). The major difference between the RMP and the NMP, is that in the former the estimates of MSY are ignored-the parameter estimates used are those for D and RY. It was shown in de la Mare (1986a) that the estimates


Fig. 1a. Trajectory of the true population, managed under the RMP using a model with A free, with deterministic absolute abundance data. Abundance overestimated.


Fig. 1c. Time series of estimates of D , from a population managed under the RMP, using a model with A free, with deterministic absolute abundance data. Abundance overestimated.
of RY from biased models could be relatively unbiased, if the rate of population change is not great. Estimates of D can be relatively unbiased (de la Mare, 1986b), so long as a yield determining parameter is allowed freedom in the fitting process, provided biased estimates of absolute abundance are not used.

## The RMP with biased estimates of absolute abundance

In de la Mare (1986b) it is shown that the estimates of depletion from fitting a population model to time series of absolute abundance data can be significantly biased if the estimates of abundance are biased. In this section, the effect of bias in the estimates of absolute abundance is examined, for the case where the model is fitted directly to the absolute abundance estimates. The trials are deterministic, with the true population having parameter set (1), and the fitted model having parameter set (2). Since these trials make relatively larger errors about initial catches, assessments start at year 10 , but with only the initial population as a free parameter. Assessments made with A as a free parameter begin at year 20. The first trial is for abundance estimates biased upwards by a factor of two; the results are shown in Figs 1a to 1d. The population trajectory (Fig. 1a) shows that the population does not converge on the target, and by year 200 appears to be approaching stability near the true protection level. This results from the bias in the estimates of depletion (Fig. 1d).


Fig. 1b. The catch history from a population managed under the RMP using a model with A free, with deterministic absolute abundance data. Abundance overestimated.


Fig. 1d. Time series of bias in estimates of $D$, from a population managed under the RMP, using a model with A free, with deterministic absolute abundance data. Abundance overestimated.


Fig. 2a. Trajectory of the true population, managed under the RMP, using a model with A free, with deterministic absolute abundance data. Abundance underestimated.


Fig. 2c. Time series of estimates of $D$, from a population managed under the RMP, using a model with A free, with deterministic absolute abundance data. Abundance underestimated.

Figs 2 a to 2 d show the results from a second trial based on absolute abundance estimates which are biased downwards by a factor of two. This case also shows a failure for the population to converge on the target. An interesting feature is that the estimates of depletion (Fig. 2c) switch to a different value just after year 160. The resultant population trajectory (Fig. 2a) shows the population initially tending to stabilise slightly above the target level, but the catches exceed MSY after the flip in depletion estimates.

Overall, the results suggest that the RMP will fail to achieve the management objectives if the filter used is a population model fitted to biased estimates of absolute abundance. Since it is not usually possible to determine if the abundance estimates are unbiased, it may be preferable to treat the absolute abundance estimates as an index of relative abundance. This modification is explored in the next section.

Treating the absolute abundance estimates as an index of relative abundance
In order to treat the absolute abundance series as an index of relative abundance, it is necessary to estimate an extra 'nuisance' parameter, the bias in the population estimates;


Fig. 2b. The catch history from a population managed under the RMP, using a model with A free, with deterministic absolute abundance data. Abundance underestimated.


Fig. 2d. Time series of bias in estimates of $D$, from a population managed under the RMP, using a model with A free, with deterministic absolute abundance data. Abundance underestimated.
this is analogous to the catchability coefficient in the catch-effort case. In general, this will give the estimation procedure the same properties as found in de la Mare (1986b) for the catch-effort case. However, a difference arises in closed loop management; there can be independence between information and control if the abundance estimates are independent of catching, for example, if the estimates are based on sightings surveys. Such is not the case if the estimates are based on marking experiments, since the size of catch plays a role in the precision of the estimates.

The time series of absolute abundance data used in these trials are normal and homoscedastic and so an ordinary least squares residual function gives maximum likelihood estimates. The residual function to be minimised is given by:

$$
\begin{equation*}
\mathrm{S}=\sum_{\mathrm{t}}\left(\mathrm{~N}_{\mathrm{t}}-\mathrm{bP} \mathrm{P}_{\mathrm{t}}\right)^{2} \tag{5}
\end{equation*}
$$

where $N_{t}$ is the abundance estimate in year $t, P_{t}$ is the estimate from the population model and $b$ is the reciprocal of the estimate of the bias in the abundance estimates. The least squares estimate of $b$, for a given vector of $P_{t}$ is given by:

$$
\begin{equation*}
\hat{b}=\sum_{t}\left(N_{t} / P_{t}\right) / \sum P_{t}{ }^{2} \tag{6}
\end{equation*}
$$



Fig. 3a. Time series of estimates of P , from a population managed under the RMP, using a model with A free, with deterministic relative abundance data.


Fig. 3c. Trajectory of the true population, managed under the RMP, using a model with A free, with deterministic relative abundance data.

The free bias parameter allows the RMP to have the same deterministic performance as the relative abundance case in de la Mare (1986a), reproduced here in Figs. 3a to 3d. The major question to confirm is whether the separation of information from control cures the problem of drifting to low levels of exploitation. One set of fifty trials is run with


Fig. 3b. Time series of estimates of $D$, from a population managed under the RMP, using a model with A free, with deterministic relative abundance data.


Fig. 3d. The catch history from a population managed under the RMP, using a model with A free, with deterministic relative abundance data.


Fig. 4. Distributions of catch, deepest depletion, true population levels and estimates of $P$ and $D$, at 30 and 100 years, for a population managed under the RMP. Absolute abundance data with a CV of 0.4 ; data treated as index of relative abundance.
absolute abundance estimates, but with the free bias parameter. The estimates of abundance are actually unbiased, so as to give approximately the same distribution of true population size in the year in which the full management procedure is first applied, as for the set of trials using relative abundance data (however, the results are not exactly comparable, because of detail differences in calculating catch limits in the period of near constant recruitmentl. Only the higher noise level is used.

The distributions of the performance parameters are shown in Fig. 4. It is clear from a comparison of this figure with Fig. 5, which shows the corresponding results using catch-effort data, that the separation of information from
control has overcome the problem of failing to realise the potential of the stocks. It is interesting to compare these results with the corresponding trials where the model is fitted directly to unbiased abundance estimates, without a free bias parameter. The two sets of trials are fully comparable, because the same set of random deviates is used in each, and there are no detail differences about the simulations other than the method used to fit the population model to the data. The results here, in terms of eventual convergence on the target level, are similar to those obtained from fitting to the absolute abundance data (Fig. 6), but with a lower variance. The absolute case has a mean true depletion, at 100 years, of $0.724(\mathrm{SD}=0.139)$.


Fig. 5. Distributions of catch, deepest depletion, true population levels and estimates of $P$ and $D$, at 30 and 100 years, for a population managed under the RMP from relative abundance data with a CV of 0.4 .


Fig. 6. Distributions of catch, deepest depletion, true population levels and estimates of $P$ and $D$, at 30 and 100 years, for a population managed under the RMP from absolute abundance data with a CV of 0.4.

The corresponding results for the relative fitting procedure is a mean true depletion of $0.780(\mathrm{SE}=0.109)$. There is also an improvement in containing the extent of accidental overexploitation; the deepest depletion in this set of trials is 0.436 ; the corresponding figure in the absolute case is 0.356 . The number of trials in which the stock was reduced to below the protection level was 5 , compared with 11 in the absolute case. However, there is an increase in the frequency of cases in which a stock was estimated to require protection; 27 in these trials, versus 16 for the absolute case.

Overall, these results reflect improved estimates of depletion in the longer term with the free bias parameter. However, this improvement is at the cost of a considerable loss of precision in the estimates of population size, plus an increase in the frequency of protecting stocks which did not require it; both the result of relaxing the constraint on fitted population size (in the absolute case, the fitted population has a similar mean abundance to that of the time series of abundance estimates). In the light of the improvement in performance, this loss of precision for P is not important; the estimates of P and A (as well as the bias) from the fitted population model could be regarded as 'nuisance' parameters, required for the reliable estimation of D.
The important point is that this variant of the management procedure will still perform reasonably well even if the absolute abundance estimates are biased, and, at least in the longer term, at no obvious loss in comparison to the unbiased absolute abundance case. Absolute abundance estimates could be biased for a variety of reasons; for example, methodological problems or failure to correctly identify the geographic or temporal distribution of stocks. Given that it is unlikely that an assumption that such estimates are unbiased can be verified, it seems that they should all be treated as indices of relative abundance. The important principle which is demonstrated by these trials is that reliable management requires independence between the acquisition of information about the stock and the resultant catches.

## Preliminary consideration of measures for improving the stability of catch limits

In de la Mare (1986a) it was found that a certainty equivalence management procedure leads to considerable year to year variability in catch limits. It is likely that such variability would not be acceptable to an industry, particularly one which exploits only one or two stocks. However, an industry which has a number of stocks available for exploitation may be sufficiently stable overall, so long as the assessments for individual stocks are not highly correlated. The variability in catch limits could be reduced, either by tuning the parameters of the control law, or by delaying the implementation of certainty equivalence control until a given amount of data were available. This topic will not be examined in detail here, but some examples will be given which indicate some of the factors to consider in designing such schemes.

The first examples are single realisations from a scheme in which the implementation of certainty equivalent control is delayed. The true population has parameter set (1), the fitted model has parameter set (2). Absolute abundance data with an approximate CV of 0.2 are used, the fitted model has $P, A$ and bias as free parameters. Initial catch limits are set conservatively, at $0.5 \%$ of the estimate of initial abundance. Two delays in implementing


Fig. 7. A single realisation of the RMP, with a 50 year delay in implementation, and a conservative start.


Fig. 8. A single realisation of the RMP, with a 100 year delay in implementation of certainty equivalence control.


Fig. 9. A single realisation of the RMP, with an early implementation of certainty equivalence control.


Fig. 10. A single realisation of the RMP, with a 50 year delay in implementation, with variation in the initial catch regime.
certainty equivalence are considered; 50 and 100 years. The resultant population trajectories and catch histories are shown in Figs 7 and 8 respectively. In both cases, the catch histories show that this approach is not particularly successful. Comparing these results with a set in which a less cautious start is made, given in Fig. 9, shows that the early implementation of certainty equivalence leads to a more rapid stabilisation of catch limits. This suggests that the larger initial catches, coupled with the higher variability resulting from the early implementation of control, is more informative about the populations dynamics. Thus, a conservative start based on constant catch limits may not lead to any significant improvement in the stability of catch limits.

An alternate strategy would be to set the average catch limit conservatively, but to vary the catch limit from year to year. This type of strategy is indicated by the studies of the 'dual control problem' (e.g. Smith and Walters, 1981). Figs 10 and 11 show two realisations of the RMP in which the catch limits are set periodically, the first at $0.75 \%$ of the initial estimate of abundance for ten years, then followed by ten years at $0.25 \%$. These two trials are fully comparable with those shown in Figs 7 and 8. The results suggest that the variable catch policy is more informative about the population's dynamics. Nonetheless, in terms of stabilising the catch limits, the results are not particularly encouraging.

A further possibility, is to start conservatively, but with a revised control law. For example, this could involve applying the management control law to the average catch, perhaps with less weight being given to catches further back in time. Such a policy would be less variable than using the estimate of RY, but it could also increase the


Fig. 11. A single realisation of the RMP, with a 100 year delay in implementation, with variation in the initial catch regime.


Fig. 12. A single realisation of a modified RMP using catch limits derived from the average catch, with implementation delayed to year 50 .
length of time required for the management procedure to stabilise the stock at the target level. This latter effect would arise because averaging the catches leads to some lag; it takes time for altered catches to have much effect on the average catch. A single realisation for this procedure is given in Fig. 12; the procedure has been quite successful in producing relatively stable catch limits. However, it is not possible to say from this one trial what the general properties of the procedure would be, particularly with respect to achieving the conservation objectives for a stock. The important point is that a more steady industry can be achieved, but this has a cost in terms of more conservative levels of exploitation in the short term.

## A FURTHER MODIFICATION TO TAKE PARAMETER UNCERTAINTY INTO ACCOUNT

It has long been recognised that there is a need for some method for taking uncertainty in parameter estimates into account when formulating management advice for whale stocks (e.g. Tillman and Chapman, 1979; Allen, 1979; 1980). In this section, a possible scheme is developed on intuitive grounds. The scheme is a modification to the RMP with explicit feedback, although the principle could be applied to other management schemes. The RMP uses two uncertain parameter estimates, namely RY and D. In principle, the catch limits could be discounted in some way to contain the risk of overexploitation, by considering the variance of one or both of these parameters. However, the RY is not a suitable candidate for discounting. If approximately unbiased estimates of the RY are available, discounting will lead to less than the RY being caught on
average, and so the population will move towards the unexploited population size. Hence, the RY will become smaller. Such a policy would result in the catch limits ultimately being reduced to zero.

An intuitively appealing scheme is to move the target level further away from the protection level, by an amount which depends upon the uncertainty in the estimate of $D$. The basis of the scheme presented here is to set the target level at a point such that the confidence region for the estimate of depletion only overlaps the protection level at a given confidence value. The expression for the discounted target level ( $\mathrm{T}^{\prime}$ ) can be written as follows:

$$
\begin{align*}
\mathrm{T}^{\prime} & =\mathrm{T}, & & \mathrm{X} \leqslant \mathrm{~T}-\mathrm{Q} \\
& =\mathrm{Q}+\mathrm{X}, & & \mathrm{X}>\mathrm{T}-\mathrm{Q}
\end{align*}
$$

where T is the notional target level, Q is the protection level and $X$ is the distance from the estimate of $D$ to its lower confidence bound. The rationale is that, when operating at the target level, the possibility that the real stock is depleted to below the protection level can be excluded to a given degree of confidence. The value of $X$ depends on the level of risk chosen which would allow catching to continue even though the real stock might be below the protection level. This interpretation treats confidence levels associated with estimates of $D$ as degrees of belief about the true value of $D$.

Examples of the effect of expression (7) on the management control law are shown in Fig. 13, for various levels of $\hat{\mathrm{S}}_{\mathrm{D}}$, assuming the $\hat{\mathrm{D}}$ is normally distributed and X is set at 1.64 SEs , to give a one-tail confidence interval at the 0.95 level. It can be seen that, above the threshold level for X , the discount not only shifts the target level upwards,


Fig. 13. Examples of the control law when modified to take uncertainty in estimates of depletion into account, for various sizes of standard error.


Fig. 14. A contour on the surface of the residual function for a joint estimate of $P$ and $A$, also showing depletion isopleths, and the locus of the point which has the minimum sum of squares for given estimates of depletion.
but also reduces the slope of the multiplier for the RY. This is an appealing feature; when the uncertainty about the estimate of D is high, the more slowly a stock apparently above the target level will be driven towards it. When the level of uncertainty is such that the confidence region for the estimates of $D$ for the stock, even at the unexploited level, does not exclude the protection level to the required level of confidence, on average, the catch limits will be less than the RY. Thus, if such levels of uncertainty were to persist, catch limits would ultimately be set at zero. It should be borne in mind that $\mathrm{T}^{\prime}$ is dynamic (and also stochastic), that is, the target level will change as uncertainty about depletion is reduced; nonetheless, $T^{\prime}$ has a lower limit of T. Another appealing feature is that, given estimates of X below the threshold level, there is no discount for uncertainty; the scheme operates as the revised scheme of the preceding section.

## Estimating confidence intervals for the depletion

The implementation of the scheme requires a method of estimating confidence intervals for $\hat{D}$. This can be done from an analysis of the residual function for a given estimate. However, this is complex because $\hat{D}$ is a derived parameter from the joint estimates of $P$ and $A$, for the given catch history. Confidence intervals, based on asymptotic theory, can be found using likelihood ratios formed from the sums of squares (Bard, 1974). The critical value for a sum of squares residual function is given by:

$$
\begin{equation*}
S=S_{o} 1+\left[(v / n-v) \cdot F_{v, n-v, \alpha}\right] \tag{8}
\end{equation*}
$$

where $S_{o}$ is the value of the residual function at the minimum, $\mathrm{F}_{\mathrm{n}}, \alpha$ is the critical value from an F distribution, at the significance level, with degrees of freedom $n$, the total number of degrees of freedom, and $v$, the number of parameters estimated. Nuisance parameters need not be counted in the degrees of freedom associated with the number of parameters estimated; the values of the residual function are those associated with the joint marginal distribution of the parameters of interest (Cox and Hinkley, 1974, pp.321-3).

Fig. 14 illustrates the procedure for finding the values of the residual function associated with the marginal distribution of the estimated value of D . The figure shows
the joint estimate of $P$ and $A$, and an arbitrary residual contour from a single simulation of fitting a population to absolute abundance data, but treated as an index of relative abundance. It also shows depletion isopleths, that is, loci on the P , A plane which give the same level of depletion, for the given series of catches. The sum of squares associated with a given degree of depletion for the marginal distribution of the estimate of D is the minimum value along the appropriate depletion isopleth, and corresponds to a unique point in the P , A plane. The locus of these points is also shown. Fig. 15 shows the value of the residual function, versus depletion, along this locus, along with the one-tail $95 \%$ confidence level. The analysis gives a


Fig. 15. Values of the residual function associated with the marginal distribution of the estimate of depletion.
one-tail lower $95 \%$ confidence bound for $D$ of approximately 0.53 . Since the estimate of D was found to be 0.78 , this result means that the target level in the modified management procedure would be set at 0.8 .

There are methods for obtaining exact confidence regions in non-linear regression (Williams, 1962; Halperin, 1963; Kirkwood, 1981) but these methods are cumbersome and are not examined here.

Using a jackknife procedure to estimate the confidence bound
The analysis of the preceding subsection is too difficult to undertake for simulation tests of the modified RMP. The trials in the following section require around one thousand estimates of the lower confidence bound for D. Therefore, an estimator is required which is more amenable to incorporation into a computer simulation program. The candidate examined here is the jackknife, extensively reviewed by Miller (1974). The jackknife technique is a resampling scheme which involves recalculating estimates from subsets of the data. The method can remove bias of order $1 / \mathrm{n}$ (where n is the number of degrees of freedom) from an estimate, and also can give an estimate of the standard error. However, it can give poor estimates if outliers are present in the data. The jackknife estimate of a parameter is obtained by splitting $n$ data into $g$ subgroups each size $h$. The jackknife estimate is given by:

$$
\begin{equation*}
\mathrm{Y}=(1 / \mathrm{g}) \sum_{\mathrm{i}=1}^{\mathrm{g}} \mathrm{Y}_{\mathrm{i}} \tag{9}
\end{equation*}
$$

where

$$
\begin{equation*}
\tilde{Y}_{i}=g \hat{Y}-(g-1) \hat{\mathrm{Y}}_{-i} \tag{10}
\end{equation*}
$$

and $\hat{Y}$ denotes the parameter estimate from the full data set, $Y_{-i}$ denotes the parameter estimate when a subset $i$ is deleted from the data. An estimate of the standard error of Y is given by:

$$
\begin{equation*}
\tilde{s}_{Y}\left[\sum_{i=1}^{g}\left(Y_{i}-Y\right)^{2}\right] / \mathrm{g}-1 \tag{11}
\end{equation*}
$$

The following statistic has an asymptotic t distribution, wit 1 g - 1 degrees of freedom:

$$
\begin{equation*}
\mathfrak{t}=\left[\sqrt{ } \mathrm{g}\left(\tilde{\mathrm{Y}}_{\mathrm{i}}-\mathrm{Y}\right)\right] / \tilde{\mathbf{s}} \tag{12}
\end{equation*}
$$

and forms a pivotal statistic for robust interval estimation (Miller, 1974). In the work presented in this paper, $h$ is unity, that is, the observations are deleted one at a time in the jackknife procedure.

In order to give an indication as to whether the jackknife procedure could be used with the modified RMP, a set of 100 simulation trials was run. The estimates are obtained by fitting a population model with parameter set (2) to 30 years of absolute abundance estimates with an approximate CV of 0.2 , derived from a population trajectory for a true population with parameter set (1). The fitted parameters are $\mathrm{P}, \mathrm{A}$, and a free bias parameter. The HI catch history, (de la Mare, 1986b) is employed. The true initial population size is 10,000 and the true degree of depletion is 0.705 . These parameters are chosen because the results of these trials will then be applicable to the simulations of the modified RMP given in the next subsection. The trials here are not aimed at a general validation of the jackknife with this type of nonlinear regression procedure.

Fig. 16 gives the distributions of both the jackknifed and ordinary estimates of $P$ and $D$, along with the jackknife


Fig. 16. Distributions of estimates of $P$ and $D$, compared with corresponding jackknifed estimates, and jackknife estimates of their standard errors.
estimates of their SDs. The results show that the jackknife estimates of P are unreliable. This is not surprising given the highly skewed distribution of the raw estimates, and the finite probability of failure to obtain an estimate. Nonetheless, the results for the estimates of D and its SD are much better behaved. This is probably a reflection of the better behaviour of the estimates of depletion, which are usually bounded even though the population estimates are unbounded. The mean of the jackknife estimates of the SD is 0.117 , close to the SD of the ordinary estimates of 0.106 . However, the distribution of the jackknife estimates of the SD indicates that these estimates may not be particularly precise. Overall, the results suggest that the jackknife estimate of the confidence interval for $D$ will be adequate for use in simulation trials of the modified management procedure.

Another potential method for estimating the lower confidence bound, related to the jackknife, is the bootstrap procedure (Efron, 1979). In the situation here, the method would involve multiple replications of the estimates, using new data sets which are created by adding randomly selected sets (with replacement) from the residuals to the estimated population trajectory. This method can give better results than the ordinary jackknife used here (Hinkley and Wei, 1984) but requires more computation (Efron and Gong, 1983) and for this latter reason is not examined here.


Fig. 17. Distributions of catch, deepest depletion, true population levels and estimates of $P$ and $D$, at 30 and 100 years, for a population managed under the modified RMP. Absolute abundance data with a CV of 0.2 , and intersample period of 4 years; data treated as index of relative abundance, and using jackknife estimates for the uncertainty in $D$.


Fig. 18. Distributions of catch, deepest depletion, true population levels and estimates of $P$ and $D$, at 30 and 100 years, for a population managed under the RMP. Absolute abundance data with a CV of 0.2 , and intersample period of 4 years; data treated as index of relative abundance.

## A simulation test of the modified RMP

One set of 50 triais is run, with absolute abundance estimates made every four years, with an approximate CV of 0.2 . The true population has parameter set (1), and the fitted model has parameter set (2). The fitted parameters are $P, A$ and $a$ bias parameter. The amount of testing which can be done in this paper is limited by the amount of computation required for the simulation; applying the jackknife procedure leads to a severalfold increase in the number of estimates required (and these are all found by a direct minimisation search, each requiring 50 to 100 population simulations). Hence, it has not been possible to undertake extensive trials of the procedure.

The jackknife confidence bounds are obtained from a t distribution, and hence, in these trials, the control law can be rewritten:

$$
\begin{array}{rlrl}
\mathrm{T}^{\prime} & =\mathrm{T}, & \tilde{\mathrm{~s}}_{\mathrm{D}} \leq(\mathrm{T}-\mathrm{Q}) / \zeta_{\mathrm{g}} \\
& =\mathrm{Q}+\zeta_{\mathrm{g}} \tilde{\mathrm{~s}}_{\mathrm{D}}, \tilde{\mathrm{~s}}_{\mathrm{D}}>(\mathrm{T}-\mathrm{Q}) / \zeta_{\mathrm{g}} \tag{13}
\end{array}
$$

where $\mathrm{s}_{\mathrm{D}}$ is the jackknife estimate of the standard error in the estimate of $\mathrm{D}, \zeta_{\mathrm{g}}$ sets the level of confidence region overlap. The magnitude of $\zeta_{\mathrm{g}}$ depends on g , the available degrees of freedom. In these trials, the confidence level chosen for the overlap with the protection level is 0.95 .
The distributions of the results are shown in Fig. 17. Fig. 18 shows a fully comparable set of trials from the RMP,
without the uncertainty modification. The results show that the modification has improved the performance of the management procedure with respect to reducing the probability of reducing a stock to below the protection level; although, on this dimension the trials have not turned out to be a very searching test. No stock was reduced below the protection level under the modified regime, compared with three instances for unmodified case. The deepest depletion was 0.576 , compared with 0.499 for the unmodified scheme. An improvement also occurs in the frequency with which stocks were estimated to require protection; 14 for the modified procedure, versus 18 for the unmodified version. The overall result in terms of mean level of exploitation is not greatly different for the two sets of trials. The mean depletion, after 100 years, in the modified case is 0.825 ( $\mathrm{SD}=0.081$ ); the corresponding result for the unmodified version is 0.780 ( $\mathrm{SD}=0.105$ ). Calculating the sustainable yield at the two different mean values of depletion gives 73.7 for the modified procedure, compared with 80.1 for the unmodified procedure. Thus, the expected annual cost of the uncertainty discount is about $8 \%$ at the 100 year point.

The modification which takes uncertainty into account does appear to lead to an improvement in the performance of the management procedure in reducing the probability of reducing a stock to below the protection level, and at a cost which seems reasonable. However, given that there is a cost; changing the amount of effort put into population surveys would change both the overall cost of management, and also the opportunity cost to the fishery. This indicates that there will be some optimum level of surveying which balances these two costs. The search for this optimum is not considered in this paper. On the basis of one trial it is not possible to say whether this modified revised management regime will perform well under all circumstances, nonetheless the results are encouraging.

## CONCLUSION

The RMP was tested in de la Mare (1986a) with unbiased estimates of absolute abundance, where it was found to give modestly successful performance in maintaining an exploited population near a target level, but with a moderate risk of the stock declining below the protection level from time to time. The trials in this paper show that if management is based on biased absolute abundance estimates, the revised procedure fails to stabilise the population at the target level. When the abundance estimates are biased upwards, the procedure stabilises the population below the target level. When the abundance estimates are biased downwards, the procedure initially tends to stabilise the population above the target level, but in the longer term, the estimated depletion becomes unstable, such that the population is driven to below the target level.

It is shown that treating the abundance estimates as an index of relative abundance, by estimating an additional free parameter, overcomes this problem, but at the cost of a loss of precision in the estimates of initial population size. The results confirm that independence between information about the status of the stocks, and catches, is sufficient to prevent the phenomenon of failing to realise the potential of the stocks, which arises with the use of catch-effort data. Overall, with this estimation method, the performance of the RMP, with respect to the accuracy in stabilising the population at the target level, seems to be no
worse than the unbiased absolute abundance case, and the precision may be somewhat better. However, there is an increase in the frequency of stocks which apparently require protection, particularly in the early years of a fishery. It is suggested that the initial population size and yield determining parameters from the fitted model may as well be regarded as nuisance parameters, whose role is to improve the accuracy of the estimates of depletion.

It is shown in de la Mare (1986b) that a certainty equivalence management procedure leads to a long period of chaotic catch limits. This raises the question as to whether such a situation could be acceptable to an industry. A certainty equivalence policy relies on the assessments being followed very closely; it may break down if advice for downward revisions in catch limits are not implemented. However, an industry exploiting a number of stocks will be more stable overall than one based on the exploitation of any individual stock. However, if a steady industry is a more important objective than one which realises the full potential of stocks in the short term, then a possible solution is to delay the implemention of the certainty equivalence control until a certain number of years of data have been collected. The relevant question then becomes: What steady catch policy contains the risk of reducing the stock below the target level before the implementation of feedback control to an acceptable level? The simple, but not necessarily optimal solution is to catch only that part of the initial standing stock which is in excess of the target level, over the initial period. This does not mean that estimates of standing stock size would be revised over that time. This general question is not pursued to any extent in this paper. However, some examples suggest that conservative catches are less informative about a population's dynamics, and so catch limits still remain variable, even with 100 years of data at the point at which certainty equivalent control is implemented. Further examples show that non-steady initial catch policies may be more informative. If the objective of stabilising catch limits is accorded a high priority, the use of the average catch in place of the replacement yield in the RMP seems to be a promising line of enquiry.
A modification to the RMP is suggested to take into account the uncertainty in the estimate of the degree of depletion. In essence, the modification is to move the target level upwards, so that if the stock were at the target level, a chosen lower confidence bound on the estimate of depletion would not overlap the protection level.

A procedure for estimating the position of the lower confidence bound, based on analysis of the surface of the residual function, is examined. However, this is found to be too complex for ready incorporation into simulations of the management procedure. The jackknife procedure for interval estimation is examined by means of simulation trials, for the specific circumstances of a test of the modified management procedure. The results suggest that the jackknife estimate of the standard error in estimated depletion is adequate for the simulation trials carried out here, although the estimates are not very precise.

The results of a single set of trials with the modified RMP are encouraging. An improvement in meeting the conservation objective for a stock is demonstrated by comparison with a set of trials which did not incorporate the $u$ certainty modification. The cost of the uncertainty modification, in terms of approximate expected loss of annual catch, is shown to be reasonable.

It should be kept in mind that all the trials in this paper are made under circumstances which are ideal, representing a flow of data about whale stocks which is more steady, and less complicated than has existed for any real whale stock. Potential problems arising from inappropriate stock boundaries, nonlinear indices of abundance, natural stochasticity or time variability in the exploited population, and so on, have not been investigated. Problems of this nature may cast some doubt on whether simple population models are adequate filters for separating management signals from noisy data; purely empirical filters may be more robust under some of the circumstances likely to be encountered in practice. This seems to be an area worthy of further study. Nonetheless, a model to be fitted to data must be capable of emulating the phenomenom observed or expected for the real system. In this sense, a population model is appropriate.

## ACKNOWLEDGMENTS

The author is grateful for reviews of this manuscript by J.R. Beddington, Imperial College, 48 Prince's Gardens, London, SW7 1LU, UK and J.G. Cooke, Department of Biology, University of York, Heslington, York, Y01 5DD, UK.

## REFERENCES

Allen, K.R. 1979. Towards an improved whale management procedure. Rep. int. Whal. Commn 29: 143-5.
Bard, Y. 1974. Nonlinear Parameter Estimation. Academic Press. New York. 341pp.
Cox, D.R. and Hinkley, D.V. 1974. Theoretical Statistics. Chapman and Hall, London. 511pp.
de la Mare, W.K. 1986a. Simulation studies on management procedures. Rep. int. Whal. Commn 36:429-50.
de la Mare, W.K. 1986b. Fitting population models to time series of abundance data. Rep. int. Whal. Commn 36: 399-418.
Efron, B. 1979. Bootstrap methods: another look at the jackknife. Ann. Stat. 7:1-26.
Efron, B. and Gong, G. 1983. A leisurely look at the bootstrap, the jackknife, and cross-validation. Am. Statistician 37: 36-48.
Goodwin, G.C. and Sin, K.S. 1984. Adaptive Filtering Prediction and Control. Prentice-Hall, Englewood Cliffs, New Jersey. 540pp.
Halperin, M. 1963. Confidence interval estimation in non-linear regression. J.R. Statist. Soc. B 24: 330-3.
Hinkley, D. and Wei, B. 1984. Improvements of jackknife confidence limit methods. Biometrika 71: 331-9.
International Whaling Commission. 1983. Report of the Scientific Committee. Rep. int. Whal. Commn 33: 43-73.
Kirkwood, G.P. 1981. Estimation of stock size using relative abundance data-A simulation study. Rep. int. Whal. Commn 31: 729-35.
Miller, R.G. 1974. The jackknife - a review. Biometrika 61: 1-15.
Smith, A.D.M. and Walters, C.J. 1981. Adaptive management of stock-recruitment systems. Can. J. Fish. Aquat. Sci. 38: 690-703.
Tillman, M.F. and Chapman, D.G. 1981. Further considerations of reasons for a more conservative approach to whale management. Rep. int. Whal. Commn 31: 601-6.
Williams, E.J. 1962. Exact fiducial limits in non-linear estimation. J. R. Statist. Soc. B 24: 125-39.

# A Feedback Strategy to Regulate Catches from a Whale Stock 

Kjartan G. Magnússon<br>Science Institute, University of Iceland, Dunhagi 3, 107 Reykjavík, Iceland<br>and<br>Gunnar Stefánsson<br>Marine Research Institute, Skúlagata 4, 101 Reykjavík, Iceland<br>Accepted November 1988


#### Abstract

A simple feedback strategy for regulating catches from a whale stock is described. This strategy modifies catch quotas on the basis of the relative slope of the regression line through past CPUE values. There are two control parameters available; g-the gain-a multiplier, which transforms the relative slope of the CPUE series to obtain the relative change in catch quota and $h$-the delay-which is the number of years used in the regression. It is demonstrated that under this harvesting strategy a unique equilibrium population level exists and it is shown how this equilibrium depends on the control parameters. The stability of this equilibrium is investigated and it is demonstrated that as long as $g$ and $h$ are not 'too high', it is indeed stable. Furthermore, the values of $g$ when oscillations in catch and population set in, are given for a particular set of biological parameters and some deterministic simulation results are given to illustrate the conclusions. The results of stochastic simulations are given. In all cases the strategy is successful in stabilising the population size and, by choosing $g$ and $h$, we can to a great extent control the population to some predetermined level, e.g. if it is low, by choosing gand/or $h$ 'sufficiently high' we can drive it to a higher level. However, since the only information which is used is the relative change in the CPUE, the strategy will not in general drive the population to the optimal level. If that is to be achieved, then some extra information-e.g. a population estimate-which enables a suitable choice of the control parameters to be made, is required.


## INTRODUCTION

During the past few years there has been a growing realisation of the need for new methods for regulating whale stocks and setting catch quotas, largely as a result of the failure of the IWC Scientific Committee to reach consensus, either on the classification of most stocks or recommending catch limits according to the 'New Management Procedure' (e.g. IWC, 1977, para. 8). So far, potential new management methods have primarily focussed on feedback controls on CPUE values or stock estimates to modify the catch quota annually (Cooke, 1989; de la Mare, 1986, 1989; Sakuramoto and Tanaka, 1989). In these papers simulations have been carried out using feedback on the slope of the CPUE series and/or on the difference between the present CPUE value (or stock estimate) and some target value which has been determined in advance. This target value should correspond to a population level close to MSY level. In Cooke (1989), regular stock estimates obtained from surveys were used in addition to the assumed or estimated values of various biological parameters such as age at recruitment and/or maturity, natural mortality, MSY level and MSY rate (MSYL and MSYR for short). Then the population level is estimated using a population model and the quota modified as a function of this estimate.

These methods require some knowledge of MSYL and population estimates from sighting surveys or from a population model (in which case knowledge of some of the biological parameters is essential). However, in general the MSYL is not known and population estimates are either not available or, if so, not reliable.

In this paper a potential management strategy is described which regulates catches on the basis of past CPUE values. This strategy will be appropriate in cases of limited information, i.e. where the CPUE series is the only information available; no estimates of present and initial
(pristine) population levels are available, no information on MSYL exists and hence the target level to which the population level should be driven is undetermined.

This management process therefore, is not concerned with trying to control the population to some predetermined level, but only to halt any decline detected from the CPUE series. Since the present population is unknown, it is possible, if a population has been driven to a low level before the management strategy is activated, that it will remain at a low level. However, if information on the degree of depletion of the stock is available, then it is possible to choose the control parameters in the feedback control in such a way that the population is driven to a higher level. The question of how this choice should be made is not considered here, but these features will become apparent in the simulation results presented below. This paper is only concerned with demonstrating that the strategy is successful in stabilising a declining population and in driving it to a higher level, for some choices of the control parameters.

The strategy only makes use of the slope of the regression line through the $\log$ CPUE values for the past $n$ years, where n is a predetermined number which can be used as one of the control variables and changed on the basis of estimates of present population level and target level (i.e. MSYL). The basic idea is to lower the annual catch quota when the slope through the past n values is negative and increase it when the slope is positive. In other words, the primary aim of the strategy is to stabilise the stock using only the relative rate of decrease/increase of the CPUE series to modify the catch quota which is the control action. This is known as derivative control, since only the derivative of the state is fed back into the control action. Sakuramoto and Tanaka (1986) use proportional plus derivative (PPD) control, i.e. using both the state and its derivative in the feedback control. Here, only derivative
control is used in view of the difficulties in obtaining reliable estimates of present population size and the target level.
In a sense we are regarding the whale stock as a black box, the input being the catch quotas and the output the CPUE values, which are taken as an index of stock size. It is assumed that stock size ( $\mathbf{P}$ ) and CPUE are related by

$$
\mathrm{CPUE}=\mathrm{qPr}, \mathrm{r}>0
$$

where q is the catchability coefficient and r is an exponent. In general $\mathrm{r} \leq 1$, (Cooke, 1985).

This strategy is in many ways similar to the one introduced by Sakuramoto and Tanaka (1989), except that no attempt is made to obtain a target level to which the population should be controlled.

In the following section we consider the continuous time version of the population model and the management strategy (i.e. using differential equations as opposed to difference equations) to show how the equilibrium population level under the described harvesting regime depends on the various control parameters. Furthermore, conditions for the stability of the equilibrium are given and for particular values of the biological parameters used in the population model, (i.e. natural mortality, MSYL and MSYR), we describe for which values of the control parameters oscillations in the catch and population level set in. Then, using a discrete time model, we give the results of some deterministic simulations for illustrative purposes. In the final section we give the results of stochastic simulations and explore possible further developments.

## EQUILIBRIUM POPULATION LEVELS AND STABILITY

In order to describe the harvesting strategy and investigate equilibrium population levels and stability, we will consider the following differential equation model (see May, 1980), which is the continuous version of the Pella-Tomlinson model used by the IWC. We assume that prior to harvesting, (i.e. for $\mathrm{t}<0$ ), the population is at carrying capacity, which is denoted by $K$, and put $p(t)=P(t) / K$, where $P(t)$ is the population size at time $t$. We then have

$$
\begin{gather*}
\mathrm{p}^{\prime}(\mathrm{t})=-\mathrm{mp}(\mathrm{t})+\mathrm{mp}(\mathrm{t}-\mathrm{T})\left\{1+\mathrm{A}\left[1-\mathrm{p}(\mathrm{t}-\mathrm{T})^{\mathrm{z}}\right]\right\}-\mathrm{C} / \mathrm{K}, \mathrm{t}>0 \\
\mathrm{p}(\mathrm{t})=1, \mathrm{t} \leq 0 \tag{2}
\end{gather*}
$$

Here the prime symbol denotes differentiation with respect to time, $A$ is a resilience parameter for the density dependence in the birth rate, $z$ is the density dependent exponent, m is the natural mortality rate and T is the time taken to attain sexual maturity. It is assumed that age at recruitment is equal to age at maturity.

The harvesting rate C is a given function of t , (usually a constant function), for an initial period $\mathrm{T}_{\mathrm{o}}$, and then C is modified in such a way that the relative change in C is proportional to the relative change of the population $h$ years ago. Thus,

$$
\begin{align*}
\mathrm{C}(\mathrm{t}) & =\mathrm{C}_{0}(\mathrm{t}), 0<\mathrm{t}<\mathrm{T}_{0}  \tag{3}\\
\mathrm{C}^{\prime}(\mathrm{t}) / \mathrm{C}(\mathrm{t}) & =\mathrm{g} \mathrm{p}^{\prime}(\mathrm{t}-\mathrm{h}) / \mathrm{p}(\mathrm{t}-\mathrm{h}), \mathrm{t} \geq \mathrm{T}_{0} \tag{4}
\end{align*}
$$

MSYL and MSYR are related to the parameters in equation (1) by:

$$
\text { MSYL }=[1 /(z+1)]^{1 / z} \text { and MSYR }=\mathrm{mAz} /(\mathrm{z}+1)
$$

Note that equation (4) can be written as

$$
\begin{equation*}
\mathrm{d}[\log \mathrm{C}(\mathrm{t})] / \mathrm{dt}=\mathrm{g} \mathrm{~d}[\log \mathrm{p}(\mathrm{t}-\mathrm{h})] / \mathrm{dt} \tag{5}
\end{equation*}
$$

We refer to the parameter $g$ as the feedback gain since the relative rate of change of the input (i.e. catch) is obtained by multiplying the relative rate of change of output (i.e. population or CPUE) by $g$. We will take $g \geq 1$. Integrating equation (5) we obtain

$$
\begin{equation*}
\mathrm{C}(\mathrm{t})=\mathrm{C}\left(\mathrm{~T}_{0}\right)\left[\mathrm{p}(\mathrm{t}-\mathrm{h}) / \mathrm{p}\left(\mathrm{~T}_{0}-\mathrm{h}\right)\right] \mathrm{g}, \mathrm{t} \geq \mathrm{T}_{0} \tag{6}
\end{equation*}
$$

Let us now consider the equilibrium stock under this harvesting strategy. The equation for the equilibrium is

$$
\begin{equation*}
\mathrm{ap}^{\mathrm{g}-1}+\mathrm{p}^{\mathrm{z}-1}=0 \tag{7}
\end{equation*}
$$

where $\mathrm{a}=\left[\mathrm{C}\left(\mathrm{T}_{\mathrm{o}}\right) / \mathrm{K}\right] /\left[\mathrm{mAp}\left(\mathrm{T}_{\mathrm{o}}-\mathrm{h}\right) \mathrm{g}\right]$.
Only one solution to this equation exists in the interval from 0 to 1 and this solution will be denoted by $p^{*}$. It is easy to see that $\mathrm{dp}^{*} / \mathrm{da}<0$. Thus the equilibrium level increases with decreasing a .

We can also show that in general $\mathrm{p}^{*}<\mathrm{p}\left(\mathrm{T}_{\mathrm{o}}-\mathrm{h}\right)$. Indeed, let us put $F(p)=\mathrm{apg}^{-1}+\mathrm{p}^{z-1}$. Then $\mathrm{F}(0)<0$ and $F(1)>0$, so a solution of (7) lies between 0 and 1 . Now

$$
\begin{aligned}
\mathrm{F}\left[\mathrm{p}\left(\mathrm{~T}_{\mathrm{o}}-\mathrm{h}\right)\right] & =\left\{\mathrm{C}\left(\mathrm{~T}_{\mathrm{o}}\right) / \mathrm{K}-\mathrm{mA}\left[1-\mathrm{p}\left(\mathrm{~T}_{\mathrm{o}}-\mathrm{h}\right)^{z}\right] \mathrm{p}\left(\mathrm{~T}_{\mathrm{o}}-\mathrm{h}\right)\right\} / \\
& {\left[\mathrm{mAp}\left(\mathrm{~T}_{\mathrm{o}}-\mathrm{h}\right)\right] } \\
& =\left\{\mathrm{C}\left(\mathrm{~T}_{\mathrm{o}}\right) / \mathrm{K}-\mathrm{f}\left[\mathrm{p}\left(\mathrm{~T}_{\mathrm{o}}-\mathrm{h}\right)\right]\right\} /\left[\mathrm{mAp}\left(\mathrm{~T}_{\mathrm{o}}-\mathrm{h}\right)\right]
\end{aligned}
$$

where $f(p)=m A\left(1-p^{z}\right) p$ is the net production rate.
Thus,
$\mathrm{F}\left[\mathrm{p}\left(\mathrm{T}_{\mathrm{o}}-\mathrm{h}\right)\right]>0$ if $\mathrm{C}\left(\mathrm{T}_{\mathrm{o}}\right) / \mathrm{K}>\mathrm{f}\left[\mathrm{p}\left(\mathrm{T}_{\mathrm{o}}-\mathrm{h}\right)\right]$ and
$\mathrm{F}\left[\mathrm{p}\left(\mathrm{T}_{\mathrm{o}}-\mathrm{h}\right)\right]<0$ if $\mathrm{C}\left(\mathrm{T}_{\mathrm{o}}\right) / \mathrm{K}<\mathrm{f}\left[\mathrm{p}\left(\mathrm{T}_{\mathrm{o}}-\mathrm{h}\right)\right]$,
that is, $\mathrm{p}^{*}<\mathrm{p}\left(\mathrm{T}_{0}-\mathrm{h}\right)$ in the case where the point [ $\left.\mathrm{p}\left(\mathrm{T}_{\mathrm{o}}-\mathrm{h}\right), \mathrm{C}\left(\mathrm{T}_{\mathrm{o}}\right) / \mathrm{K}\right]$ lies above the net production rate curve and $\mathrm{p}^{*}>\mathrm{p}\left(\mathrm{T}_{\mathrm{o}}-\mathrm{h}\right)$ if it lies below. The second case would only apply in cases where a substantial reduction in catch rate has taken place before the feedback control is 'switched on'. If the catch rate is constant in the initial period, then $\mathrm{p}^{*}<\mathrm{p}\left(\mathrm{T}_{\mathrm{o}}-\mathrm{h}\right)$.

This is also clear from the intersection between the harvesting rate curve, which always passes through the point $\left[\mathrm{p}\left(\mathrm{T}_{\mathrm{o}}-\mathrm{h}\right), \mathrm{C}\left(\mathrm{T}_{\mathrm{o}}\right) / \mathrm{K}\right]$, and the production rate curve as shown in Figs 1A and 1B for the values $p\left(T_{0}-h\right)=0.7$, MSYL $=0.6$, MSYR $=2 \%$ and $g=3$. If $\left[p\left(T_{o}-h\right), C\left(T_{o}\right) / K\right]$ is above the net production rate curve, then the point of intersection, whose first coordinate is the equilibrium $\mathrm{p}^{*}$, is to the left of $p\left(T_{o}-h\right)$, (Fig. 1A), and to the right if $\left[p\left(T_{o}-h\right), C\left(T_{o}\right) / K\right]$ is below the curve (Fig. 1B).
There are four external variables which will affect the position of the equilibrium:
(i) $\mathrm{C}\left(\mathrm{T}_{\mathrm{o}}\right)$, the catch rate at the end of the initial period (usually the catch is taken to be constant throughout - if the catch rate in this period is very variable then $C\left(T_{o}\right)$ can be taken as the average over the previous years);
(ii) $\mathrm{T}_{\mathrm{o}}$, the length of the initial catch period;
(iii) g , the feedback gain;.
(iv) $h$, the delay in the management process.

Furthermore, the equilibrium depends on $p\left(T_{0}-h\right)$ but this is determined by $C_{o}, T_{o}$ and $h$.

Let us now investigate how the equilibrium $p^{*}$ depends on the control parameters above.
Since $a=\left[C\left(T_{o}\right) / K\right] /\left[m A p\left(T_{o}-h\right) g\right]$, we have $\log a=\log$ $\left[C\left(T_{o}\right) / K\right]-\log (m A)-g \log p\left(T_{o}-h\right)$. Thus, $\delta(\log a) / \delta h=-g p^{\prime}-$


Fig. 1. Production and harvesting rate curves (see text).
$\left(\mathrm{T}_{\mathrm{o}}-\mathrm{h}\right)(-1) / \mathrm{p}\left(\mathrm{T}_{\mathrm{o}}-\mathrm{h}\right)<0$, since $\mathrm{p}^{\prime}\left(\mathrm{T}_{\mathrm{o}}-\mathrm{h}\right)<0 \quad$ (population decreases in the initial period). Hence we conclude that the coefficient a decreases with increasing $h$ and thus $\mathrm{p}^{*}$ increases with h . In other words, the longer the delay in the management process, the higher the equilibrium level.

We can show that $\mathrm{dp}^{*} / \mathrm{dg}=\left[\log p\left(\mathrm{~T}_{\mathrm{o}}-\mathrm{h}\right)-\log \mathrm{p}^{*}\right]\left(\mathrm{p}^{*}\right) \mathrm{g}-1 /$ $\left[(\mathrm{g}-1)\left(\mathrm{p}^{*}\right)^{\mathrm{g}-2}+\mathrm{z}\left(\mathrm{p}^{*}\right)^{z-1 / a}\right]>0$, since $\mathrm{p}^{*}<\mathrm{p}\left(\mathrm{T}_{\mathrm{o}}-\mathrm{h}\right)$, and hence it follows that $\mathrm{p}^{*}$ increases with increasing feedback gain g .
Similarly, $\delta(\log a) / \delta T_{o}=-g p^{\prime}\left(T_{o}-h\right) / p\left(T_{o}-h\right)>0$, and hence $\mathrm{p}^{*}$ decreases with increasing $\mathrm{T}_{\mathrm{o}}$.
In practice management of a whale stock does not start with the stock being at equilibrium, so in general there is no control over $\mathrm{C}_{\mathrm{o}}$ or $\mathrm{T}_{\mathrm{o}}$. Thus $\mathrm{p}\left(\mathrm{T}_{\mathrm{o}}\right.$-h), (determined by the catch history) is what, in addition to the two control variables $g$ and $h$, determines the equilibrium.
Fig. 2 shows how $\mathrm{p}^{*}$ depends on the gain g for the particular parameter values MSYL $=0.60, \mathrm{C}_{\mathrm{o}} / \mathrm{K}=3 \%$ and $\mathrm{p}\left(\mathrm{T}_{\mathrm{o}}-\mathrm{h}\right)=0.70 ; \mathrm{p}^{*}$ as a function of g is shown for MSYR 1. 2, 3 and $4 \%$.

Fig. 3 shows how $\mathrm{p}^{*}$ depends on the delay $h$ for the same values of MSYL and $\mathrm{C}_{\mathrm{o}} / \mathrm{K}$ as in Fig. 2, when g equals 3 and $p\left(T_{o}\right)$-the population when the control strategy is activated-equals 0.50 . The reason the equilibrium decreases with increasing MSYR for large $h$ is that for a high MSYR, the population decreases more slowly; $p\left(T_{o}-h\right)$ is thus lower since the population at $T_{0}$ is 0.50 in all cases. Although in these calculations, $\mathrm{p}\left(\mathrm{T}_{\mathrm{o}}-\mathrm{h}\right)$ is found by using the discrete time model described below, the difference between the solutions of the continuous and the discrete model is negligible.

Let us now consider the stability of this equilibrium. It has been shown that the equilibrium for a differential-delay equation is locally asymptotically stable, provided the


Fig. 2. Population size versus gain for various MSY rates.


Fig. 3. Population size versus delay, $h$ (see text).
linearised system is stable (Cushing, 1975). Putting $p(t)=p^{*}+x(t)$ in equation (1) and dropping second order terms we obtain

$$
\begin{array}{r}
\mathrm{x}^{\prime}(\mathrm{t})=-\mathrm{mx}(\mathrm{t})+\mathrm{m}\left[1-\mathrm{m}\left(\mathrm{p}^{*}\right)^{\mathrm{z}} \mathrm{z}+\mathrm{A}\left(1-\left(\mathrm{p}^{*}\right)^{\mathrm{z}}\right)\right] \mathrm{x}(\mathrm{t}-\mathrm{T})- \\
\\
\end{array} \mathrm{Amag}^{\left(\mathrm{p}^{*}\right) \mathrm{g}-1 \mathrm{x}(\mathrm{t}-\mathrm{h}),}
$$

and setting $x(t)=x(0) e^{t}$, we get the equation for $\lambda$

$$
\begin{equation*}
\lambda=-\mathrm{m}-\mathrm{m}\left[\mathrm{~A}(\mathrm{z}+1)\left(\mathrm{p}^{*}\right)^{z-1}-\mathrm{A}\right] \mathrm{e}^{-\lambda \mathrm{T}}-\mathrm{gmA}\left[1-\left(\mathrm{p}^{*}\right)^{z}\right] \mathrm{e}^{-\lambda \mathrm{h}} . \tag{8}
\end{equation*}
$$

If all solutions to (8) have negative real part, then it follows that the linearised system is stable (e.g. May, 1981). In order to simplify (8), we will restrict ourselves to the case where $T=h$. Then (8) becomes

$$
\begin{equation*}
\lambda=-\mathrm{m}-\mathrm{m}\left[\mathrm{~A}(\mathrm{z}+1)\left(\mathrm{p}^{*}\right)^{\left.\mathrm{z}-1-\mathrm{A}+\mathrm{gA}-\mathrm{gA}\left(\mathrm{p}^{*}\right)^{\mathrm{z}}\right] \mathrm{e}^{-\lambda \mathrm{T}},{ }^{2} \mathrm{t}}\right. \tag{9}
\end{equation*}
$$

Let us now denote $\left[\mathrm{A}(\mathrm{z}+1)\left(\mathrm{p}^{*}\right)^{\mathrm{z}}-1-\mathrm{A}+\mathrm{gA}-\mathrm{gA}\left(\mathrm{p}^{*}\right)^{\mathrm{z}}\right]$ by D . Then it can be shown (see May, 1981) that all solutions to (9) have negative real parts if

$$
\begin{equation*}
\mathrm{mT}<(\pi-\arccos (1 / \mathrm{D})) / \sqrt{ }\left(\mathrm{D}^{2}-1\right) \tag{10}
\end{equation*}
$$

This condition can be checked for each set of parameters, but there are too many of them for us to be able to obtain a more 'illuminating' stability condition for general values of the parameters involved, especially since $\mathrm{p}^{*}$ depends on $p\left(T_{0}-h\right)$. However, for specified values of some of the parameters we can see for what value of $g$ oscillations set in, since it is known (May, 1981), that as the stability boundary in (10) is crossed, there occurs a bifurcation from the stable limit point $\mathrm{p}^{*}$, to a stable limit cycle with period

$$
\begin{equation*}
\tau=[2 \pi /(\pi-\arccos (1 / \mathrm{D}))] \mathrm{T} . \tag{11}
\end{equation*}
$$



Fig. 4. Population size versus gain (see text).

Let us now assume that MSYL is at 0.60 , (i.e. $z=2.39$ ), $\mathrm{m}=0.05$ years $^{-1}$ and $\mathrm{T}=7$ years, in accordance with the recommendations of the Workshop on Management (IWC, 1988). We further assume that $\mathrm{p}\left(\mathrm{T}_{\mathrm{o}}\right)=0.50$ and that $\mathrm{C}_{\mathrm{o}} / \mathrm{K}=3 \%$.

Then, first solving $m T=(\pi-\arccos (1 / D)) / V\left(D^{2}-1\right)$, we obtain $\mathrm{D}=5.15$ and hence the period $\tau$, when oscillations set in, is 24.9 years. Then we solve for $g$ in the equation $\mathrm{D}=5.15$, noting that $\mathrm{p}^{*}$ depends on g , (see Fig.4). Table 1 shows how the solution, $g_{0}$-the value of $g$ when bifurcation to the stable limit cycle occurs-depends on MSYR. Note that the equilibrium $P$ is locally asymptotically stable for $g<g_{o}$. Furthermore, as $g$ is increased beyond $g_{o}$ and sustained oscillations set in, the amplitude of the oscillations increases with increasing g .

Table 1
The relationship between MSYR, $g_{o}$ and $p^{*}$ (see text)

| MSYR: $1 \%$ | $2 \%$ | $3 \%$ | $4 \%$ |  | $1 \%$ | $2 \%$ | $3 \%$ | $4 \%$ |  |
| :--- | ---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $g_{0}$ | 30.1 | 14.7 | 9.6 | 7.2 | $p^{*}$ | 0.586 | 0.560 | 0.528 | 0.511 |

For illustrative purposes a few population and catch trajectories resulting from the use of the management strategy described above were calculated.

The population model used was the discrete analogue to the differential equation (1):

$$
\begin{equation*}
\mathrm{P}(\mathrm{t}+1)=[\mathrm{P}(\mathrm{t})-\mathrm{C}(\mathrm{t})] \mathrm{S}+(1-\mathrm{S})\left\{1+\mathrm{A}\left[1-(\mathrm{P}(\mathrm{t}-\mathrm{T}) / \mathrm{K})^{\mathrm{z}}\right]\right\} \mathrm{P}(\mathrm{t}-\mathrm{T}) \tag{12}
\end{equation*}
$$

In accordance with the recommendations of the IWC Workshop on Management (IWC, 1988), the following parameter values were used:

$$
\begin{align*}
& \mathrm{T}=7 \text { (age at maturity) }  \tag{13}\\
& \mathrm{S}=\exp (-0.05) \text { (annual survival rate) } \\
& \mathrm{A}=0.6916 \text { corresponding to MSYR of } 2.5 \% \\
& \mathrm{~K}=10,000 \text { (initial population) } \\
& \mathrm{z}=2.39 \text { corresponding to MSYL of } 0.60
\end{align*}
$$

These values give an MSY of 150 animals. Note that the relationship between A and MSYR is now MSYR $=[\exp (m)-1] A z /(z+1)$ since this is the discrete case. The feedback strategy corresponding to (4) is:

$$
\begin{equation*}
C(t+1)=C(t)\left[1+g b_{n}(t)\right] \tag{14}
\end{equation*}
$$



Fig. 5. Population trajectories for various values of gain when $h=7$.
where $b_{n}(t)$ is the slope of the regression line of $\log$ CPUE against $t$, using the past $n$ CPUE values. The reason log CPUE is used rather than CPUE, is that only the relative rate of change of the population is measured by the CPUE and furthermore, a declining population usually declines exponentially rather than linearly. Thus we can regard the slope of the regression line as an estimate of $\mathrm{rd}(\log \mathrm{P}) / \mathrm{dt}$, (or $\mathrm{r}(\mathrm{dP} / \mathrm{dt}) / \mathrm{P}$ ) if the relationship between CPUE and P is CPUE $=\mathrm{qP}^{\mathrm{r}}, \mathrm{r}>0$; it is possible that $\mathrm{r}<1$, (Cooke, 1985).

Thus, since the last n values of the CPUE series are used in the regression, we can regard $b_{n}(t)$ as an estimate of rdlog $\mathrm{P}(\mathrm{t}-\mathrm{h}) / \mathrm{dt}$, where $\mathrm{h}=(\mathrm{n}-1) / 2$; i.e. the relative rate of change of $\mathrm{P},(\mathrm{n}-1) / 2$ years ago.

In order to demonstrate that (14) corresponds to (4), let us integrate equation (4) from $t$ to $t+1$ to obtain $\mathrm{C}(\mathrm{t}+1)=\mathrm{C}(\mathrm{t})[\mathrm{P}(\mathrm{t}+1-\mathrm{h}) / \mathrm{P}(\mathrm{t}-\mathrm{h})] \mathrm{g}$. Then, integrating this equation from $t=i$ to $t=i+1$ (i.e. integrating the catch rate to get total catch over the year), and assuming a pulse fishery, i.e.

$$
C(t)=\Sigma C_{i} \delta\left(t-t_{i}\right)
$$

where $\delta$ is the delta function-this equation means that the catch in year $i C_{i}$, is caught at the time instant $t_{i}$-we get

$$
\begin{aligned}
\mathrm{C}_{\mathrm{i}+1} & =\mathrm{C}_{\mathrm{i}}\left[\mathrm{P}\left(\mathrm{t}_{\mathrm{i}}+1-\mathrm{h}\right) / \mathrm{P}\left(\mathrm{t}_{\mathrm{i}}-\mathrm{h}\right)\right] \mathrm{g} \\
& =\mathrm{C}_{\mathrm{i}}\left[1+\left(\mathrm{P}\left(\mathrm{t}_{\mathrm{i}}+1-\mathrm{h}\right)-\mathrm{P}\left(\mathrm{t}_{\mathrm{i}}-\mathrm{h}\right)\right) / \mathrm{P}\left(\mathrm{t}_{\mathrm{i}}-\mathrm{h}\right)\right] \mathrm{g} \\
& =\mathrm{C}_{\mathrm{i}}\left[1+(\Delta \mathrm{P} / \mathrm{P})_{\mathrm{i}-\mathrm{h}}\right] \mathrm{g} \\
& \cong \mathrm{C}_{\mathrm{i}}\left[1+\mathrm{g}(\Delta \mathrm{P} / \mathrm{P})_{\mathrm{i}-\mathrm{h}}\right], \text { for }(\Delta \mathrm{P} / \mathrm{P})_{\mathrm{i}-\mathrm{h}} \text { small. }
\end{aligned}
$$

If $\mathrm{P}(\mathrm{t})$ declines exponentially, i.e. $\mathrm{P}(\mathrm{t})=\mathrm{P}(0) \exp (-\mathrm{bt})$, then the relative change in P at time $\mathrm{t}, \Delta \mathrm{P} / \mathrm{P}$ equals $\exp (-b)-1=-b+0\left(b^{2}\right) \cong-b$ for small $b$. Since $\log$ $P\left(t_{i}+1-h\right)-\log P\left(t_{i}-h\right)=-b$, it is clear that equation (14) in the discrete model corresponds to equation (4) in the continuous model.

In the simulation it is assumed that CPUE is linearly related to stock size. However, if $r$ is different from 1, then the same population trajectory will be obtained by multiplying the $g$ used for the linear case by $1 / \mathrm{r}$. The simulations in this section are deterministic in the sense that there is no error term in the relationship between CPUE and P. We take the initial catch to be 300 per year, i.e. $3 \%$ of initial stock size, and activate the feedback control when the population is down to $30 \%$ of initial level.

The results are given in Figs 5 and 6 for $h=7$, i.e. $n$, the number of years used in the regression is 15 . It is clear that increasing the gain gives a higher equilibrium value but also leads to increasing oscillations. Increasing $g$ still further would lead to sustained oscillations, and it can be


Fig. 6. Catch levels over time for the simulation given in Fig. 5.
seen from the figures that the period for $g$ equal to 10 and 14 is about 25 years as expected. It is also clear that it is not possible to drive the population beyond approximately 0.40 before sustained oscillations set in. In order to achieve that objective, it is necessary to increase $h$.
The average yearly catch over $50,100,150$ and 200 years for different values of g is shown in Table 2. It is apparent that the advantage of maintaining a higher population level does not manifest itself until after 150 years.

## Table 2

Average yearly catches over various time periods for different different values of $\mathbf{g}$

|  | Gain |  |  |  |
| ---: | :--- | :--- | :--- | :--- |
| Year | 2 | 6 | 10 | 14 |
| 50 | 300 | 300 | 300 | 300 |
| 100 | 215.8 | 208.6 | 206.7 | 201.9 |
| 150 | 168.8 | 178.6 | 181.0 | 183.2 |
| 200 | 144.7 | 163.9 | 167.4 | 168.1 |

The corresponding average catch figures for the strategy of bringing the population immediately down to MSYL and then taking MSY annually are; 230, 190, 177 and 170. It is only after 200 years that this strategy produces any advantage.

The corresponding population and catch trajectories for $\mathrm{h}=12$, i.e. $\mathrm{n}=25$, are shown in Figs 7 and 8. There oscillations set in for smaller values of $g$-which is to be expected since we are increasing the delay-and for $g=10$ the oscillations are far too great for the strategy to be a feasible one. The period in the oscillations is around 50 years. It should be obvious from these simulations that if the population is at a very low level when the management strategy is activated, then it is not possible to drive it to a high level. This is not unexpected since $p\left(T_{o}-h\right)$ restricts the values of $P$ obtainable. In order to obtain a $\mathrm{p}^{*}$ value of 0.60 , say, it is necessary to use some additional means, apart from the feedback strategy. For example, a substantial reduction in catch could be imposed for a number of years before the strategy is activated. Such measures are of course dependent on the existence of some knowledge of the actual state of the stock.

## STOCHASTIC SIMULATIONS

Some simulations have been conducted, roughly along the lines indicated by the IWC Comprehensive Assessment


Fig. 7. Population trajectories for various values of gain when $\mathrm{h}=12$.


Fig. 8. Catch levels over time for the simulation given in Fig. 7.

Workshop on Management (IWC, 1988), and are reported here although the full preliminary screening is not complete.

Simulations were conducted using lognormally distributed errors in CPUE. The random numbers were generated using the Tausworthe Feedback Shift Register Method, as described in Kennedy and Gentle (1980) for uniforms, followed by a Box-Mueller transform to obtain normals.

The following parameters have been used:
(1) lognormal density of CPUE errors, with CV=0.4 ( $\sigma=0.385$ in the corresponding normal density)-some 'deterministic simulations' were also conducted;
(2) the Pella-Tomlinson population model as in equation (12);
(3) $\mathrm{A}=0.2766$ (i.e. corresponding to $\mathrm{MSYR}=1 \%$ ), $\mathrm{T}=7$, $\mathrm{z}=2.39, \mathrm{~K}=10,000, \mathrm{~S}=\exp (-0.05)$ as described in (13).
The method described in the paper allows flexibility in the following parameters:
$\mathrm{C}_{\mathrm{o}}=$ initial constant catches $(=100,300,400$ and 500 here)
$\mathrm{T}_{\mathrm{o}}=$ number of initial constant catch years (20 and 30 here)
$\mathrm{n}=$ number of years used in the regression ( $=\mathrm{T}_{\mathrm{o}}$ here)
$\mathrm{g}=$ the feedback gain ( $=2,4$ and 6 here)
Note that choosing $\mathrm{C}_{\mathrm{o}}$ and $\mathrm{T}_{\mathrm{o}}$ is not an essential part of the strategy. Simulations are conducted here for different values of these parameters in order to bring the population down to different levels when the strategy is activated.

In Figs 9-29 (p. 177 ff .), each figure describes a set of simulations for one particular combination of $C_{o}, T_{o}=n$ and $g$.
The legend contains information about the parameters used. Here N is the number of simulated time series, g is the feedback gain used in the procedure, C is the initial constant catch level, R is the number of years used in the regression of $\log C P U E$ against time and $E$ denotes the error distribution, either 0 for deterministic version or $L$ for lognormal distribution.

For each combination, the deterministic version (no error in CPUE) of the trend in catches and stock over 100 years is shown as (a) and (b).

A sample of five catch trends and stock trends using the lognormal errors is also shown-(e and h). This is done by simply generating five time series and plotting them against time.

The remainder of each figure comprises histograms of generated outcomes of the simulations. For example, (d) is the frequency distribution of generated cumulative catches after 30 years, obtained by repeatedly generating a time series and recording the cumulative catch after 30 years. After generating e.g. 1,000 time series, one has recorded 1,000 cumulative catch values, which are represented by the histogram. Note that the histograms reflect absolute frequency, i.e. the number of simulated cumulative catches observed in the given interval; (g) gives the same type of histogram, after 100 years. A histogram of the frequency distribution of the lowest population point is given as (c). In (f), a histogram of generated population size after 30 years is shown; the same histogram, but after 100 years, is given in (i).

The simulation results can be analysed in a number of different ways, depending on the primary objective. If management of the resource without endangering the species is the primary objective, one can examine the simulation results with an emphasis on the cumulative catches ( d and g ) and the minimum population (c). It should be noted that in the cases considered here, the procedure seems very robust as far as stock extinction is concerned, although in a number of cases the stock remains somewhat away from the optimal level (i.e. MSYL).

To understand the workings of the procedure, however, it is necessary to place emphasis on the catch and stock trends for the deterministic case, (a) and (b) and for the stochastic case, (e) and (f).

When the initial catch is only 100 animals per year, the stock remains at a high level and the procedure is unable to increase the catches to appropriate levels, for all values of $g$ and n .

At the other extreme, when $\mathrm{C}_{\mathrm{o}}=500$, the stock goes down rapidly during the initial stage, but after the constant-catch years, the procedure cuts down the catch almost immediately, to almost zero, as is seen in part (e) of the figures for $\mathrm{C}_{\mathrm{o}}=500$. It is also seen in part (h) that the procedure does lead to some recovery of the stock and in part (c) one notes that the stock never drops to zero. Finally, for $\mathrm{C}_{\mathrm{o}}=500$ we note that only in the case of large feedback, $g=6$, does the procedure tend to get the stock back to the ratio $\mathrm{P} / \mathrm{K}=60 \%$ at which the model gives MSY (see (i)).

For $C_{o}=300$ and 400 , the effects of varying $g$ in a stochastic environment are clear in that more feedback (larger $g$-values) tends to do more to revive a stock that has been reduced below MSYL. However, increasing the
g -values also incorporates more of the variability of the CPUE directly into the catch quotas. This is reflected in the increasing variability when, for example, (g) is viewed for the different $g$-values. For $C_{0}=300$ and $n=20$, the total catch values with $g=2$ have a much tighter (and right-shifted) distribution than for $g=6$. For $C_{o}=300, g=4$ and 6 will tend to make the final population close to the desired $60 \%$, but when $\mathrm{C}_{\mathrm{o}}=400$, one needs $\mathrm{g}=6$ (see (i)).

Simulations were also carried out for MSYR $=4 \%$ ( $\mathrm{A}=1.1066$ ). The results are shown in Figs $30-33$ and confirm the conclusions reached above. The population is stabilised at a higher level than for MSYR $=1 \%$ and there is a tendency for oscillations to occur.

## DISCUSSION

The simulations described above demonstrate that in all cases the strategy is successful in stabilising stock. In no case does extinction occur. However, since the only objective is to halt a decline in the stock size, the strategy will not in general maximise the utilisation of the stock. This is of course to be expected, since we only observe-and hence employ-the relative rate of change of the stock size. Thus, it is only this rate we can control. If one wants to control the actual state of the stock to a specified level, observations thereof are required.

A stock which is well above MSYL at $\mathrm{T}_{\mathrm{o}}$ will remain underexploited, i.e. well above MSYL, unless a low feedback gain is used, in which case the stock will stabilise at a level somewhat lower than that at $\mathrm{T}_{\mathrm{o}}$.

On the other hand, a stock which is severely depleted will remain so, if low gain values are employed. The stock can be made to recover somewhat by using higher gain values and more years in the regression, but there is a limit to this recovery because of oscillations which set in when these control parameters are increased too much.

Thus, it is clear that the strategy, although successful in stabilising a stock for most choices of the control parameters $g$ and $n$, requires additional information if it is to succeed in driving the stock close to some 'desired' level. This information, which would have to be an estimate of the deviation of present stock level from the desired level, would be utilised to select appropriate values of $g$ and $n$. These parameters would be selected by using the equation for the equilibrium $\mathrm{p}^{*}$, i.e. equation (7).

Furthermore, if the stock is at a very high or very low level at $\mathrm{T}_{\mathrm{o}}$, it is desirable to use some other method to take the stock to some interval around MSYL, (say 40-80\% of initial level), and then start the feedback control action, with an appropriate choice of the control parameters on the basis of a population estimate, to ensure that the stock stays close to MSYL, (or to the 'desired' level). This 'other method', in the case of high $\mathrm{P}\left(\mathrm{T}_{\mathrm{o}}\right)$, could involve, for example, maintaining constant catches or using very low g -values, ( $\mathrm{g}=1$ would ensure that the stock decreased still further), until the population is in the prescribed interval. In the case of low $\mathrm{P}\left(\mathrm{T}_{\mathrm{o}}\right)$, a sharp reduction in catch quotas could be prescribed initially and then the feedback control employed on the basis of these reduced catches. These questions will not be explored further in this paper.

However, one should not place to much emphasis on steering the stock to the so called MSYL which may not be a 'real' biological concept. or, if so, may be time-varying. In any case, the yield will only drop by less than $2 \%$ when the stock size deviates $10 \%$ from MSYL. In absolute numbers the change is almost negligible in stocks with low MSYR.

Thus, the control objective should be to ensure that the stock is in no danger of extinction and that a 'reasonable' yield is obtained. Both these objectives could be met by maintaining the stock size in an interval, say $50-70 \%$ of initial level. This interval could be better determined on the basis of some knowledge of the production curve for each particular stock. This information might have to be obtained by means of some sort of probing strategy; for example by setting high catch quotas initially and driving the stock to a low level and then decreasing the catch sufficiently to ensure that the stock increases. Nonlinear discriminant analysis could be then used on the (CPUE, catch)-data to obtain some information on the shape of the yield curve. It should be noted though, that the dynamics of a whale stock are very sluggish, and therefore such a procedure would require a longer timespan than might be feasible. These questions require further attention and will not be pursued further in this paper.

## ACKNOWLEDGEMENTS

This work was supported by the Programme for Whale Research at the Marine Research Institute, Reykjavík, Iceland. We are grateful to Mr. Th. Gunnlaugsson and Prof. Th. Helgason, University of Iceland, Reykjavík for reading the manuscript and for many useful comments.

## REFERENCES

Cooke, J. G. 1985. On the relationship between catch per unit effort and whale abundance. Rep. int. Whal. Commn 35: 511-19.
Cooke, J.G. 1989. Simulation studies of two whale stock management procedures. (Published in this volume.)
Cushing, J.M. 1975. An operator equation and bounded solutions of integrodifferential systems. SIAM J. Math. Anal. 6(3): 433-45.
De la Mare, W.K. 1986. Simulation studies on management procedures. Rep. int. Whal. Commn 36: 429-50.
De la Mare, W.K. 1989. Further simulation studies on management procedures. (Published in this volume.)
International Whaling Commission. 1977. International Convention for the Regulation of Whaling, 1946. Schedule (As amended by the Commission at its 29th Meeting, June 1977). International Whaling Commission, Cambridge. 15pp.
International Whaling Commission. 1988. Report of the Comprehensive Assessment Workshop on Management Reykjavík, March 1987, Rep. int. Whal. Commn 38: 163-70.
Kennedy, W.J. and Gentle, J.E. 1980. Statistical Computing. Marcel Dekker, Inc., New York. 591pp.
May, R.M. 1980. Mathematical models in whaling and fisheries management. pp. 1-64 In: AMS, Lectures on Mathematics in the Life Sciences, Vol. 13 .
May, R.M. 1981. Non-linear phenomena in ecology and epidemiology. Ann. N.Y. Acad. 357: 267-81.
Sukuramoto, K. and Tanaka, S. 1989. A simulation study on management of whale stocks considering feedback systems. (Published in this volume)


Fig. 9. Initial catches: 100, for 20 years. Feedback gain: 2. Years in regression: 20. $\mathrm{A}=.2766$.


Fig. 10. Initial catches: 100 , for 20 years. Feedback gain: 4. Years in regression: $20 . \mathrm{A}=.2766$.

|  |  |  |
| :---: | :---: | :---: |
|  |  |  |
|  |  |  |

Fig. 11. Initial catches: 100 , for 20 years. Feedback gain: 6. Years in regression: 20. $A=.2766$.

|  |  |  |
| :---: | :---: | :---: |
|  |  |  |
|  |  |  |

Fig. 12. Initial catches: 100 , for 30 years. Feedback gain: 2. Years in regression: 30 . $A=.2766$.


Fig. 13. Initial catches: 100, for 30 years. Feedback gain: 4. Years in regression: 30. $\mathrm{A}=.2766$.


Fig. 14. Initial catches: 100 , for 30 years. Feedback gain: 6. Years in regression: $30 . \mathrm{A}=.2766$.


Fig. 15. Initial catches: 300 , for 20 years. Feedback gain: 2 . Years in regression: 20. $A=.2766$.

REP. INT. WHAL. COMMN (SPECIAL ISSUE II)


Fig. 16. Initial catches: 300 , for 20 years. Feedback gain: 4 . Years in regression: 20. $\mathrm{A}=.2766$.


Fig. 17. Initial catches: 300 , for 20 years. Feedback gain: 6. Years in regression: 20. $A=.2766$.


Fig. 18. Initial catches: 300 , for 30 years. Feedback gain: 2 . Years in regression: $30 \mathrm{~A}=.2766$.
(a) Catch trend (1 simulations, error: none)

Fig. 19. Initial catches: 300, for 30 years. Feedback gain: 4. Years in regression: $30 . A=.2766$.

|  |  |  |
| :---: | :---: | :---: |
|  |  |  |
|  |  |  |

Fig. 20. Initial catches: 300 , for 30 years. Feedback gain: 6 . Years in regression: $30 . \mathrm{A}=.2766$.


Fig. 21. Initial catches: 400 , for 20 years. Feedback gain: 2. Years in regression: 20. $A=.2766$.


Fig. 22. Initial catches: 400 , for 20 years. Feedback gain: 4 . Years in regression: 20. $\mathrm{A}=.2766$.


Fig. 23. Initial catches: 400, for 20 years. Feedback gain: 6. Years in regression: 20. $A=.2766$.

|  |  |  |
| :---: | :---: | :---: |
|  |  |  |
|  |  |  |

Fig. 24. Initial catches: 400 , for 30 years. Feedback gain: 2 . Years in regression: $30 . \mathrm{A}=.2766$.


Fig. 25. Initial catches: 400, for 30 years. Feedback gain: 4. Years in regression: 30. $\mathrm{A}=.2766$.

|  |  | (c) Minimum population (1000 simulations, error: lognormal) |
| :---: | :---: | :---: |
|  |  |  |
|  |  |  |

Fig. 26. Initial catches: 400 , for 30 years. Feedback gain: 6. Years in regression: 30. $\mathrm{A}=.2766$.

|  | (b) Population trend (1 simulations, error: none) |  |
| :---: | :---: | :---: |
|  |  | (f) Population year 30 ( 1000 simulations, error: lognormal) |
|  |  |  |

Fig. 27. Initial catches: 500, for 20 years. Feedback gain: 2 . Years in regression: 20. $A=.2766$.

|  | (b) Population trend (l simulations, error: none) | (c) Minimum population (1000 simulations, error: lognormal) |
| :---: | :---: | :---: |
|  |  |  |
|  | (h) Population trend (5 simulations, error: lognormal) | (i) Population year 100 ( 1000 simulations, error: lognormal) |

Fig. 28. Initial catches: 500, for 20 years. Feedback gain: 4. Years in regression: 20. $\mathrm{A}=.2766$.


Fig. 29. Initial catches: 500, for 20 years. Feedback gain: 6. Years in regression: 20. $\mathrm{A}=.2766$.


Fig. 30. Initial catches: 300 , for 20 years. Feedback gain: 4. Years in regression: 20. $A=1.1066$.


Fig. 31. Initial catches: 300 , for 30 years. Feedback gain: 2. Years in regression: 30 . $A=1.1066$.


Fig. 32. Initial catches: 400 , for 20 years. Feedback gain: 4 . Years in regression: $20 . \mathrm{A}=1.1066$.


Fig. 33. Initial catches: 500, for 20 years. Feedback gain: 4. Years in regression: 20. $\mathrm{A}=1.1066$.

# Results of First Stage Screening Trials for a Proposed Whale Stock Management Procedure 

A.E. Punt and D.S. Butterworth<br>Department of Applied Mathematics, University of Cape Town, Rondebosch 7700, South Africa

Accepted September 1989


#### Abstract

The background to the development of the 'Punt-Butterworth' proposed whale stock management procedure is sketched, and its three components are described in detail: the estimation procedure; the catch control law; and the additional restrictions imposed on catch limit variations from one year to the next. A Bayes-like estimator is used, incorporating a prior distribution for MSY\%, with a population model being fitted to both CPUE and absolute abundance data. The catch control law is a variant of the New Management Procedure control law, which adds a probing component at high population levels. Catch limits are generally restricted to change by less than $20 \%$ between years. Full results of the first stage screening trials are given. These are compared to results where the proposed control law is replaced by that of the New Management Procedure, and where the 'protection' level is increased from 20\% to $40 \%$ of the pre-exploitation population level.


## BACKGROUND

The primary purpose of this document is to detail the 'Punt-Butterworth' whale stock management procedure that was proposed in Punt and Butterworth (1989b) and to list the results of first stage screening trials (IWC, 1989) for that procedure. First, however, a brief summary of earlier work on this problem by the authors, which was reported in Punt and Butterworth (1988; 1989a; 1989b) is given to illustrate in particular the reasons why certain previous lines of investigation were discontinued in favour of the current approach.

There were three major considerations underlying the authors' initial attempts to develop a management procedure (Punt and Butterworth, 1988). The first was an attempt to merge empirical and population-model-fitting approaches to the problem. For the initial years of management, an empirical approach using only the slope of the CPUE series to adjust catch limits was applied. This approach was very similar to an earlier version of the 'Magnusson-Stefansson' procedure (Magnusson and Stefansson, 1989). However, once sufficient data were available to provide an adequate fit to a population model, the parameters of this fit were used instead of the CPUE slope to set catch limits. These latter catch limits were set with the intention of moving the stock to a target level in the vicinity of conventional assumptions for MSYL. In principle, this procedure could solve the particular problem of an approach based only on the slope of the CPUE series, which may underutilise the stock by stabilising it either far above or far below MSYL.
The second consideration was to speed the computation of the model fits. The population model (IWC, 1988; 1989) used to generate the data for the tests agreed is an observation-error model, i.e. the population dynamics are deterministic, and errors occur only in the observations of absolute abundance from sighting surveys and of relative abundance from CPUE data. Estimation procedures which assume an observation-error structure to fit such models to data are computationally intensive because they require a non-linear minimisation. Instead a process-error estimator was used, which assumes that observations are exact measures of abundance, but that errors are present in the
model of the stock's dynamics. This allows the model to be cast in a form that is linear in certain combinations of the parameters, so that estimates for $\mathrm{A}, \mathrm{K}$ and q (resilience, pre-exploitation level and catchability respectively) are simply and quickly obtained by a ( $3 \times 3$ ) matrix inversion, and estimates of the variances of these parameters are also readily calculated. Both absolute abundance and CPUE data were used to fit the model, with the latter being smoothed to reduce the effects of the underlying observation error (eventually 15 point quadratic smoothing was used). The model fit was only accepted and used if it satisfied certain 'reasonableness' criteria. These included requiring that the estimate of the resilience parameter, A, fell in a range corresponding to an MSY\% between $0 \%$ and $5 \%$, and that deterministic dynamics projections from the estimated pre-exploitation population level, K, were sufficiently close to observations at the start and at the end of the available data series.

The third consideration was to promote inter-annual catch stability. To this end, the New Management Procedure (NMP) catch control law used was adjusted so that the range of population levels over which catch limits changed from zero to 0.9 MSY was widened from [ 0.54 K , 0.60 K ] to $[0.2 \mathrm{~K}, 0.7 \mathrm{~K}]$. In addition, restrictions were placed on the extent to which catch limits could change from one year to the next.

The attempted merging of empirical and model-fitting approaches was not successful. It was found that catch limits based on the slope of the CPUE series tended to produce inadequate contrast in population size to allow sufficiently precise estimation of the parameters of the full population model at a later stage. This led to difficulties in moving populations back towards MSYL in rehabilitation scenarios (population size $\mathrm{P}_{0}=0.3 \mathrm{~K}$ at the start of management) if no CPUE data were available prior to that time.

Thus, work reported in Punt and Butterworth (1989a) reverted to an approach based entirely on fitting population models. If there were too few data to fit the 3-parameter linear process-error estimator described above, or the resultant fit failed to satisfy the 'reasonableness' criteria, a hierarchy of simpler estimation
procedures was then attempted. These culminated in a 'bottom line' observation-error estimator, which fixed $\mathrm{MSY} \%=1 \%$ in the fitting process and used only absolute abundance data; the resultant population level estimates were then substituted into the catch control law with MSY\% fixed at $1.5 \%$. Results for rehabilitation scenarios were found to be relatively insensitive to whether or not CPUE data were available prior to the commencement of management.

While this approach incorporating a multi-level estimation procedure performed satisfactorarily for rehabilitation scenarios, and development scenarios ( $\mathrm{P}_{0}=\mathrm{K}$ at the start of management) with MSY\% $=1 \%$, substantial underutilisation occurred for development scenarios with MSY\% $=4 \%$. Considerable improvement in this regard was achieved by introducing a probing component into the catch control law to increase the data contrast in development scenarios, and so to allow the model parameters to be estimated more precisely. This involved increasing the value of MSY\% in the NMP-like catch control law when the 'bottom line' estimator was used and the population size was estimated to be above 0.7 K . The improvement was, however, at the expense of increased variation in catch limits during the early years of exploitation.

A disadvantage of the procedure developed in Punt and Butterworth (1989a) was the considerable complexity of the hierarchy of estimation procedures involved. Attempts were made to simplify these in Punt and Butterworth (1989b), but the simplifications investigated all gave rise to the problem of under-utilisation of stocks with MSY\% = $4 \%$.

Punt and Butterworth (1989b) compared the performance of the complex procedure of Punt and Butterworth (1989a) (with a few new modifications) to that of the proposed procedure described below, for the first stage screening trials. While average total catches and average final and lowest depletions were very similar for the two approaches, the standard deviations of these statistics and measures of inter-annual catch variability were much greater for the complex procedure. The results of this comparison, together with considerations of greater simplicity, led to the procedure below being preferred despite an approximately three-fold increase in the computation time required to carry out trials.

## THE PROPOSED PROCEDURE

The procedure consists of three parts: the method used to estimate parameters when fitting the population model to the data (the 'estimator'); the catch control law; and additional restrictions imposed on the inter-annual variation of catch limits. These are detailed below in turn.

## The Estimator

The population dynamics model used to fit the data has the same structure as the operating model used to generate the data for the first stage screening trials (IWC, 1988) viz:
$P_{i+1}=\left(P_{i}-C_{i}\right) e^{-M}+\left(1-e^{-M}\right) P_{i-t_{m}}$

$$
\begin{equation*}
\left\{1+\mathrm{A}\left[1-\left(\mathrm{P}_{\mathrm{i}-\mathrm{t}_{\mathrm{m}}} / \mathrm{K}\right)^{\mathrm{z}}\right]\right\} \tag{1}
\end{equation*}
$$

where $P_{i}$ is the number of whales at the beginning of year $i$, $C_{i}$ is the number of whales caught during a pulse fishing season at the start of year $i$,

A is the MSY\% determining resilience parameter, M is the instantaneous natural mortality rate, K is the pre-exploitation population size, $t_{m}$ is the age at maturity, and $z$ is a parameter which determines the ratio of MSYL to $K$. The parameters $\mathrm{M}, \mathrm{t}_{\mathrm{m}}$ and z are fixed at $\mathrm{M}=0.05 \mathrm{yr}^{-1}, \mathrm{t}_{\mathrm{m}}$ $=7 \mathrm{yrs}$ and $\mathrm{z}=2.39$ which corresponds to MSYL $=0.6 \mathrm{~K}$. (These choices are the same as the values used in the underlying model which generates the data for the first stage screening trials.) The estimator provides estimates of parameters A (or equivalently MSY\%) and K, as well as a time series of population $\left(\mathrm{P}_{\mathrm{i}}\right)$ estimates.

The estimator uses both absolute abundance estimates $\left(\mathrm{P}_{\mathrm{i}} \mathrm{s}\right)$ and CPUE data $\left[(\mathrm{C} / \mathrm{E})_{\mathrm{i}}\right]$, and assumes an observation-error structure. To compensate for the effects of possible non-linearity in the CPUE-abundance relationship, the estimator assumes that:

$$
\begin{equation*}
(\mathrm{C} / \mathrm{E})_{\mathrm{i}}=\mathrm{q}\left(\mathrm{P}_{\mathrm{i}}\right)^{0.75} \tag{2}
\end{equation*}
$$

where the 'catchability' parameter, q , has also to be estimated when fitting the model to the data. The functional to be minimised in fitting the model to determine the catch limit for year $n$ is then:

$$
\begin{align*}
\mathrm{SS}=\sum_{\mathrm{i}=0}^{\mathrm{n}-1}\left\{\ln (\mathrm{C} / \mathrm{E})_{\mathrm{i}}-\ln \left(\mathrm{q} \hat{\mathrm{P}}_{\mathrm{i}}^{0.75}\right)\right\}^{2} & +2 \sum_{\mathrm{j}}^{\mathrm{j}=0} \mathrm{n} \\
& \left\{\ln \left(\hat{\mathrm{P}}_{\mathrm{j}}^{\mathrm{s}}\right)-\ln \left(\hat{\mathrm{P}}_{\mathrm{j}}\right)\right\}^{2} \tag{3}
\end{align*}
$$

Note that CPUE data are available for each year from the commencement of management $(i=0)$ to year $n-1$, unless there was a zero catch in the year concerned. The $\Sigma^{\prime}$ covers only those years (every fifth year commencing in year 0 ) for which absolute abundance estimates are available from sighting surveys. It is assumed that such estimates are obtained before the season commences, so that an estimate for year $n$ could be used in the process of setting a catch limit for that year. The factor 2 in equation (3) is fixed, and gives an absolute abundance estimate twice the weight of a CPUE data point in the fitting procedure. Given values of A (or MSY\%) and K, the corresponding estimate of q can be obtained directly by differentiating equation (3) with respect to q , so that the minimisation involves a search over only two parameters, which are chosen to be $K$ and MSY\%.

The application of this fitting procedure without any additional constraints may result in very high MSY\% estimates, whose use could lead to rapid stock depletion. This is particularly the case during the early years of management, when information is minimal (often only one or two absolute abundance estimates and a few CPUE points) so that MSY\% is very poorly determined. Thus, the estimator proposed attempts to incorporate prior information to constrain the estimates of MSY\%. First, MSY\% is restricted to the range [ $0.5 \%, 5 \%$ ]. Further, to reduce inter-annual variation in estimates of MSY\% and resultantly in catch limits, a prior distribution for MSY\% is introduced. The prior has a maximum value at MSY\% = $1.5 \%$. If this maximum is $h$, then the prior is defined as:
$\mathrm{p}(\mathrm{MSY} \%)= \begin{cases}0 & \mathrm{MSY} \%<0.5 \% \\ 1+(\mathrm{h}-1)(\mathrm{MSY} \%-0.5) & 0.5 \% \leq \mathrm{MSY} \% \leq 1.5 \% \\ 1+(\mathrm{h}-1)(5-\mathrm{MSY}) / 3.5 & 1.5 \% \leq \mathrm{MSY} \% \leq 5 \% \\ 0 & \text { MSY\% }>5 \%\end{cases}$
Tests of performance for different values of $h$ using the first stage screening trials resulted in the choice of $h=1.5$ to obtain an appropriate balance between utilisation of the resource and increasing the risk of depletion. The distribution corresponding to this choice is shown in Fig. 1.


Fig. 1. The prior distribution for MSY\%. The points corresponding to the discrete values of MSY\% allowed in the estimation procedure are indicated.

To speed numerical computations, only integral and half-integral values of MSY\% in the range $[0.5 \%, 5 \%$ ] are considered. For a fixed MSY\%, minimisation of SS in equation (3) requires a search over the $K$ parameter alone. The resultant likelihood is proportional to $e^{-S s}$. This is multiplied by $\mathrm{p}(\mathrm{MSY} \%)$ to provide a posterior distribution for MSY\%, and the maximum of this distribution over the discrete set of MSY\% values considered provides MŜY\% (as well as the associated $\hat{\mathrm{q}}, \hat{\mathrm{K}}$ and $\left\{\hat{\mathrm{P}}_{\mathrm{i}}: \mathrm{i}=0,1, \ldots, \mathrm{n}\right\}$ ). Note that under the assumption of unimodality of the posterior distribution, a search over all the allowed values of MSY\% is usually not necessary in each year that additional data becomes available; if the posterior probability for the previous estimate of MSY\% remains larger than that for the allowed MSY\% values immediately above and below, MŜY\% remains unchanged. This aspect makes considerable computing time savings possible.

This Bayes-like estimation approach means that all management procedures start with an estimate of MSY\% $=1.5 \%$ (see Fig. 2b), and change this (in multiples of $0.5 \%$ ) only as the accumulation of data permits improved estimation of MSY\% therefrom.

## The Catch Control Law

The basic catch control law formula is:

where MŜY $=0.006 \mathrm{M} \hat{S} Y \% \hat{K}$

$$
\mathrm{M} \tilde{\mathrm{~S}}=0.03 \hat{\mathrm{~K}}
$$

i.e. MS̃Y is the MSY estimate corresponding to an MSY\% of $5 \%$ for MSYL $=0.6 \mathrm{~K}$.

Equation (5) results in the family of harvesting algorithms which is illustrated in Figs 2a and 2b. All are zero for $P_{n} \leqslant 0.2 \mathrm{~K}$. They increase linearly to 0.9 MSY at $\mathrm{P}_{\mathrm{n}}$ $=0.7 \mathrm{~K}$, and then linearly (but with a change of slope) to $4.5 \%$ of MSYL (i.e. $2.7 \%$ of $K$ ) at $P_{n}=K$. These harvesting algorithms thus set large catch limits if the stock is assessed to be near its pre-exploitation size. Nevertheless, the maximum MSY\% estimate of $5 \%$ allowed means that even if the estimate of MSY\% is greater than the true value, the catch limit cannot be set larger than $0.027 \hat{\mathrm{~K}}$. Thus, the harvesting algorithm
provides a probing mechanism to generate the additional data contrast desirable for precise parameter estimation (as discussed above), but at the same time provides an upper bound (in a deterministic sense) on catch limits.
This catch control law will eventually stabilise the stock at a target level close to 0.7 K for the range of MSY\% values considered (see Fig. 2a).


Fig. 2a. Diagrammatic illustration of the basic catch control law component of the proposed management procedure. The solid curves are the sustainable yield as a function of population size (Pella-Tomlinson form with MSYL $=0.6 \mathrm{~K}$ ) for various MSY\%'s, and the dashed lines indicate the catch limits corresponding to various population sizes and the same set of estimated MSY\%'s.


Fig. 2b. The complete set of catch control laws of equation (5) between which the estimator chooses are shown by the dashed lines. Initially the solid line corresponding to MSY\% $=1.5 \%$ is selected, but this selection is modified to the other curves as the estimate of MSY\% changes with the acquisition of further data as time progresses.

An alternative and more conservative estimator-control law combination was tested in which the estimate of $K$ was reduced by one and by two estimated standard errors (evaluated using the information matrix method conditioned upon the current estimate of A being exact). Note that such standard errors (and consequent catch limit reductions) are large initially when few data are available, but decrease with time, so that the degree of 'conservatism' applied is a function of the extent of the information available about the stock. The performance of the more conservative procedures was much worse in MSY $\%=4 \%$
scenarios, as the stocks were left well above MSYL and much smaller total catches were achieved. In other scenarios, the gains in final population sizes were very slight. Thus the procedure which did not reduce K in this manner was preferred (i.e. use of the prior distribution for MSY\% alone seems to provide adequate protection against unintended depletion of the resource because of estimation imprecision).

## Restrictions on Catch Limit Variations

The basic catch limit given by equation (5) is modified, if necessary, to conform to the following restrictions:
$\begin{cases}\left|C_{n}-C_{n-1}\right| \leq 0.2 C_{n-1} & \text { if } \hat{P}_{n} \hat{K}>0.35 \\ C_{n}=0 & \text { if } \hat{P}_{n} \hat{K}<0.15 \\ C_{n}=0 & \text { if } \hat{P}_{n} \hat{K}<0.25 \text { and } C_{n-1}=0 \\ 0.5 C_{n-1} \leq C_{n} \leq 1.2 C_{n-1} & \text { otherwise }\end{cases}$
This equation imposes a maximum inter-annual catch limit variation of $20 \%$, but only provided the stock is estimated to be in excess of $0.35 \hat{\mathrm{~K}}$. For stock level estimates in the range $[0.15 \hat{\mathrm{~K}}, 0.35 \hat{\mathrm{~K}}]$, catch limits can be reduced by up to $50 \%$ from one year to the next. If the stock is estimated to be below $0.15 \hat{K}$, catch limits are immediately set to zero, and no further catches may take place until the stock is estimated to have recovered to $0.25 \hat{\mathrm{~K}}$ at least. Although the first stage screening trials are such as never to test this last aspect of the procedure, it was felt that such a feature should be incorporated because of the need to allow for rapid downward adjustment of catch limits when serious degrees of depletions are detected, irrespective of industrial stability considerations.

## RESULTS AND DISCUSSION OF FIRST STAGE SCREENING TRIALS

IWC (1989) requires that eight performance attributes are to be reported for each set of simulation trials. In the Tables that follow, a further attribute has been added. The lower $95 \%$ confidence limit of the distribution for the lowest level to which the stock is depleted has been reported because the distribution of this depletion level is often non-normal. This results in the mean less two standard deviations not giving a good indication of the possible extent of depletion, which is an important factor in assessing the risk associated with a proposed management procedure. The nine attributes reported in the Tables are thus the following:
Mean final depletion: the average of the ratio of the final population size after 100 years of management to the pre-exploitation population size.
$S D$ (f.d.): the standard deviation of the ratio of the final population size after 100 years of management to the pre-exploitation population size.
Mean lowest depletion: the average of the ratio of the lowest population size over the 100 years of management to the pre-exploitation population size.
$S D$ (l.d.): the standard deviation of the ratio of the lowest population size over the 100 years of management to the pre-exploitation population size.
Lower 95\% limit (l.d.): the lower 95\% confidence limit of the distribution of ratio of the lowest population size over the 100 years of management to the pre-exploitation
population size. This attribute is computed by ordering the lowest depletions from each of the 100 trials and averaging the 2 nd and 3 rd values.

Total catch: the average total catch over the 100 year management period.
$S D$ (t.c.): the standard deviation of the total catch over the 100 simulation trials.
RMS catch difference: the average over the 100 simulation trials of the root mean square inter-annual catch change over the 100 years of management.
CV catch: the average over the 100 simulation trials of the coefficient of variation of the catch over the 100 years of management.
The averages were taken using the random number generator provided by the IWC Secretariat, which was activated in an order proposed in documents circulated between the developers of the various proposed management approaches.

## Table 1

Results of first stage screening trials for the proposed whale management procedure. Scenario descriptions refer to the error structure of the operating models used to generate the data : sighting survey CV/CPUE abundance relationship [square root (S) or linear (L)]. f.d. $=$ final depletion; I.d. $=$ lowest depletion; * $=$ of mean lowest depletion; diff. $=$ difference

| Attribute | Scenarios |  |  |  |
| :---: | :---: | :---: | :---: | :---: |
|  | 0.4/S | 0.2/S | 0.4/L | 0.2/L |
| Trial no. 1, $P_{0}=K, M S Y \%=1 \%$ |  |  |  |  |
| Mean f.d. (SD) | 0.571 (0.061) | 0.575(0.041) | $0.634(0.056)$ | 0.634 (0.042) |
| Mean 1.d. (SD) | 0.560(0.055) | 0.569(0.035) | 0.604(0.039) | 0.612 (0.030) |
| Lower 95\% limit* | 0.465 | 0.506 | 0.547 | 0.568 |
| Total catch (SD) | 11008 (863) | 10982 (580) | 10102 (825) | 10129 (614) |
| RMS catch diff. | 12.3 | 9.9 | 11.5 | 9.4 |
| CV catch | 0.538 | 0.459 | 0.622 | 0.545 |
| Trial no. 2, $P_{0}=K, M S Y \%=4 \%$ |  |  |  |  |
| Mean f.d. (SD) | 0.806(0.061) | 0.807(0.044) | 0.815(0.050) | 0.814 (0.037) |
| Mean 1.d. (SD) | 0.767(0.054) | $0.788(0.038)$ | 0.776(0.048) | $0.795(0.031)$ |
| Lower 95\% limit* | 0.658 | 0.703 | 0.679 | 0.733 |
| Total catch (SD) | 16787 (2470) | 16915 (2090) | 16582 (2210) | 16710 (1853) |
| RMS catch diff. | 16.8 | 14.5 | 16.8 | 14.5 |
| CV catch | 0.320 | 0.247 | 0.315 | 0.244 |
| Trial no. 3, $P_{0}=0.3 \mathrm{~K}, \mathrm{MSY} \%=1 \%$ |  |  |  |  |
| Mean f.d. (SD) | 0.469(0.056) | 0.471 (0.048) | $0.445(0.048)$ | 0.446 (0.041) |
| Mean I.d. (SD) | 0.299(0.005) | 0.300 (0.000) | 0.300(0.004) | 0.300(0.000) |
| Lower 95\% limit* | 0.300 | 0.300 | 0.300 | 0.300 |
| Total catch (SD) | 2805 (650) | 2775 (566) | 3107 (555) | 3082 (463) |
| RMS catch diff. | 4.2 | 3.5 | 4.3 | 3.6 |
| CV catch | 0.514 | 0.387 | 0.487 | 0.362 |
| Trial no. 4, $P_{0}=0.3 \mathrm{~K}, \mathrm{MSY} \%=4 \%$ |  |  |  |  |
| Mean f.d. (SD) | 0.831 (0.056) | 0.825 (0.059) | 0.771 (0.057) | 0.755 (0.052) |
| Mean 1.d. (SD) | 0.300(0.000) | 0.300 (0.000) | 0.300(0.000) | 0.300(0.000) |
| Lower 95\% limit* | 0.300 | 0.300 | 0.300 | 0.300 |
| Total catch (SD) | 11039 (2871) | 11264 (2669) | 13706 (2486) | 14293 (2169) |
| RMS catch diff. | 12.3 | 10.8 | 13.7 | 12.1 |
| CV catch | 0.593 | 0.577 | 0.591 | 0.576 |

## The Proposed Procedure

Table 1 presents the results of the 16 stochastic simulation trials required by IWC (1989). The error structures and CPUE-abundance relationships in the Table are:

| Column | Sighting survey <br> CV | Relationship between <br> CPUE and abundance |
| :---: | :---: | :---: |
| $0.4 / \mathrm{S}$ | 0.4 | square root |
| $0.2 / \mathrm{S}$ | 0.2 | square root |
| $0.4 / \mathrm{L}$ | 0.4 | linear |
| $0.2 / \mathrm{L}$ | 0.2 | linear |

Table 2
Results of the $0.4 / \mathrm{S}$ trials for the proposed management procedure (column A) and two variants of the NMP control law without and with inter-annual catch variation restrictions (columns $B$ and $C$ respectively) (see text for full details)

| Attribute | Management procedures |  |  |
| :---: | :---: | :---: | :---: |
|  | A | B | C |
| Trial no. 1, $P_{0}=K, M S Y \%=1 \%$ |  |  |  |
| Mean final depletion (SD) | 0.571 (0.061) | $0.684(0.046)$ | 0.686(0.046) |
| Mean lowest depletion (SD) | 0.560(0.055) | 0.681 (0.045) | 0.683(0.045) |
| Lower 95\% limit (1.d.) | 0.465 | 0.596 | 0.593 |
| Total catch (SD) | 11008 (863) | 8337 (782) | 8304 (795) |
| RMS catch difference | 12.3 | 9.7 | 6.5 |
| CV catch | 0.538 | 0.229 | 0.221 |
| Trial no. 2, $P_{0}=K, M S Y \%=4 \%$ |  |  |  |
| Mean final depletion (SD) | 0.806(0.061) | 0.871 (0.059) | 0.873(0.059) |
| Mean lowest depletion (SD) | 0.767 (0.054) | 0.856(0.058) | $0.858(0.057)$ |
| Lower 95\% limit (l.d.) | 0.658 | 0.757 | 0.757 |
| Total catch (SD) | 16787 (2470) | 10765 (2159) | 10664 (2128) |
| RMS catch difference | 16.8 | 13.9 | 9.6 |
| CV catch | 0.320 | 0.283 | 0.274 |
| Trial no. 3, $P_{0}=0.3 \mathrm{~K}, \mathrm{MSY} \%=1 \%$ |  |  |  |
| Mean final depletion (SD) | 0.469 (0.056) | 0.546(0.077) | 0.529(0.076) |
| Mean lowest depletion (SD) | 0.299(0.005) | 0.299(0.006) | 0.298 (0.010) |
| Lower 95\% limit (l.d.) | 0.300 | 0.300 | 0.292 |
| Total catch (SD) | 2805 (650) | 2109 (976) | 2304 (929) |
| RMS catch difference | 4.2 | 16.1 | 9.8 |
| CV catch | 0.514 | 1.751 | 482 |
| Trial no. 4, $P_{0}=0.3 \mathrm{~K}, \mathrm{MSY} \%=4 \%$ |  |  |  |
| Mean final depletion (SD) | 0.831 (0.056) | 0.838 (0.064) | 0.830(0.059) |
| Mean lowest depletion (SD) | 0.300(0.000) | 0.300(0.000) | 0.300(0.000) |
| Lower 95\% limit (1.d.) | 0.300 | 0.300 | 0.300 |
| Total catch (SD) | 11039 (2871) | 10996 (3038) | 11007 (2973) |
| RMS catch difference | 12.3 | 36.5 | 18.3 |
| CV catch | 0.593 | 0.804 | 0.738 |

Except for the MSY\% $=4 \%$ rehabilitation ( $\mathrm{P}_{0}=0.3 \mathrm{~K}$ ) scenarios, the means of the statistics differ little (irrespective of the relationship between CPUE and abundance and the CV of the absolute abundance estimates). In the MSY\% $=4 \%$ rehabilitation scenario, a linear relationship between CPUE and abundance results in larger total catches. As would be expected, the standard deviations are larger when the sighting survey CV is 0.4 . Also as would be expected, final population sizes are (slightly) lower in development ( $\mathrm{P}_{0}=\mathrm{K}$ ) scenarios where the square root relationship between CPUE and abundance holds.

These results show that the proposed management procedure performs satisfactorily in all of the 16 first stage screening trials specified by IWC (1989) and thus that it is ready for second stage testing.

## A Procedure using the NMP Control Law

The results reported in this section address the question of whether the performance of a management procedure consisting of the New Management Procedure (NMP) control law and the observation-error estimator proposed in this paper, is substantially improved by using this estimator with the proposed catch control law (equation 5). The results of this latter management procedure for the case of a sighting survey CV of 0.4 and a square root relationship between CPUE and abundance (indicated by $0.4 / \mathrm{S}$ ) are listed again as column A in Table 2. This comparison is pertinent because the IWC Comprehensive Assessment Workshops on Management Procedures have been set up to provide a 'new' whale management
algorithm; if it can be shown that the NMP control law (with an appropriate estimation procedure) is able to perform satisfactorily, the only change that may need recommendation might be the specification of the estimation procedure to be used.

The basic limit catch algorithm for the NMP control law is:
$C_{n}= \begin{cases}0 & \text { if } \hat{\mathrm{P}}_{\mathrm{n}} \leq 0.54 \hat{\mathrm{~K}} \\ 15 \mathrm{MS} Y\left[\left(\hat{\mathrm{P}}_{\mathrm{n}} / \hat{\mathrm{K}}\right)-0.54\right] & \text { if } 0.54 \hat{\mathrm{~K}}<\hat{\mathrm{P}}_{\mathrm{n}}<0.6 \hat{\mathrm{~K}} \\ 0.9 \mathrm{MSY} & \text { if } \hat{\mathrm{P}}_{\mathrm{n}} \geq 0.6 \hat{\mathrm{~K}}\end{cases}$
This is illustrated in Fig. 3.


Fig. 3. Diagrammatic illustration of the NMP catch control law of equation (7). The solid curves are the sustainable yield as a function of population size (Pella-Tomlinson form with MSYL $=0.6 \mathrm{~K}$ ) for various MSY\%'s, and the dashed lines indicate the NMP catch limits corresponding to various population sizes and the same set of estimated MSY\%'s.

The results of an 'NMP' management procedure with this control law and the estimator proposed in this paper for the $0.4 / \mathrm{S}$ simulation trials are reported in column B of Table 2. Note that for these results, the basic catches of equation (7) are NOT modified by the inter-annual catch limitation restrictions of equation (6). The results of this 'NMP' management procedure together with the following rule used to modify the basic catch limit:

$$
\begin{equation*}
0.8 C_{n-1} \leq C_{n} \leq 1.2 C_{n-1} \quad \text { if } C_{n-1} \geq 20 \tag{8}
\end{equation*}
$$

are presented in Table 2 (column C).
Comparison of the results in Table 2 shows that there is hardly any difference between the two 'NMP' variants (columns B and C) in terms of total catches and final depletions. However, the addition of the inter-annual catch variability restriction (column C) substantially improves the performance of this 'NMP' in terms of RMS catch differences.
In most respects there is little difference between the 'constrained NMP' (column C) and the procedure proposed in this paper (column A). Catch variability is greater for the latter procedure for the development ( $\mathrm{P}_{0}=\mathrm{K}$ ) scenarios, because of the probing nature of the catch control law of equation (5). However, such variability is greater for the 'NMP' in rehabilitation ( $\mathrm{P}_{\mathbf{0}}=$ 0.3 K ) scenarios because it sets catches to zero more frequently. For the MSY\% = $1 \%$ scenarios, the choice between the procedures essentially involves what
trade-offs between slightly higher catches (the proposed procedure) and slightly lesser levels of depletion (the 'NMP') are deemed desirable.
However, the proposed procedure completely outperforms the 'NMP' in one particular scenario - the development case with MSY\% $=4 \%$ - where the proposed procedure produces a $57 \%$ improvement in total catch and leaves the population closer to the nominal target level of 0.7 K . This is essentially the result of the probing component of the proposed control law of equation (5), which the NMP control law of equation (7) lacks.

Thus, even though the performance of the NMP control law can be improved by using it with the estimator proposed in this paper and adding an inter-annual catch variability restriction, the proposed procedure is to be preferred because of its better performance for the development case with MSY\% $=4 \%$.

## A Sensitivity Test with respect to the Catch Control Law Parameters

In order to determine the sensitivity of the performance of the proposed management procedure to the value chosen for the 'protection' level [i.e. the level at which catches are set to zero by the catch control law of equation (5)], this basic control law was altered to:
$C_{n}= \begin{cases}0 & \text { if } \hat{P}_{n} \leq 0.4 \hat{K} \\ 3 M \hat{S} Y\left[\left(\hat{P}_{n} / \hat{K}\right)-0.4\right] & \text { if } 0.4 \hat{K}<\hat{P}_{n}<0.7 \hat{K} \\ 0.9 \text { MS̃Y }+3(M \tilde{S} Y-M \hat{S} Y)\left(1-\hat{P}_{n} / \hat{K}\right) & \text { if } \hat{P}_{n} \geq 0.7 \hat{K}\end{cases}$
and the inter-annual catch fluctuation restrictions of equation (6) modifying this basic control law to:
$\begin{cases}\left|C_{n}-C_{n-1}\right| \leq 0.2 C_{n-1} & \text { if } \hat{P}_{n} / \hat{K}>0.55 \\ C_{n}=0 & \text { if } \hat{P}_{n} / \hat{K}<0.35 \\ C_{n}=0 & \text { if } \hat{P}_{n} / \hat{K}<0.45 \text { and } C_{n-1}=0 \\ 0.5 C_{n-1} \leq C_{n} \leq 1.2 C_{n-1} & \text { otherwise }\end{cases}$

This means that the 'protection' level was raised from 20\% to $40 \%$ of K and the catch 'prohibition' level (i.e. the level at which catches are actually set to zero) was raised from $15 \%$ to $35 \%$ of K. The control law above provides a greater degree of protection by reducing catch limits more sharply when the population level is estimated to be low.

Table 3 (column B) presents the results for this harvesting algorithm combined with the proposed estimator for the $0.4 / \mathrm{S}$ trials, and is to be compared with the results for the proposed management procedure which are given in column A. The performance of the two procedures in the development ( $\mathrm{P}_{0}=\mathrm{K}$ ) scenarios is scarcely different, as the population rarely drops low enough to be influenced substantially by the choice of the 'protection' level.

In the rehabilitation $\left(\mathrm{P}_{0}=0.3 \mathrm{~K}\right)$ scenarios, the results generally show more variation for the control law of equation (9). For the MSY\% $=4 \%$ scenario, the final population levels are similar but there is a loss of total catch; this occurs because the lesser rate of accumulation of CPUE data means that the correction of the MSY\% estimate to a higher value is only achieved at a later time, after the population has passed through the range where it is most productive. As would be expected, there is a trade-off between the total catch taken and the final population size for the MSY $\%=1 \%$ case; the choice in this instance is essentially a matter which depends on the precise objectives of the Commission.

## Table 3

Results of the $0.4 / \mathrm{S}$ trials for the proposed management procedure (column A) and a variant which sets the 'protection' level at $40 \%$ instead of $20 \%$ of K (column B)

| Attribute | Management procedures |  |
| :---: | :---: | :---: |
|  | A | B |
| Trial no. 1, $P_{0}=K, M S Y \%=1 \%$ |  |  |
| Mean final depletion (SD) | 0.571 (0.061) | $0.580(0.063)$ |
| Mean lowest depletion (SD) | 0.560(0.055) | 0.563 (0.055) |
| Lower 95\% limit (1.d.) | 0.465 | 0.466 |
| Total catch (SD) | 11008 (863) | 10886 (878) |
| RMS catch difference | 12.3 | 12.5 |
| CV catch | 0.538 | 0.557 |
| Trial no. 2, $P_{0}=K, M S Y \%=4 \%$ |  |  |
| Mean final depletion (SD) | 0.806(0.061) | $0.808(0.061)$ |
| Mean lowest depletion (SD) | 0.767 (0.054) | 0.767 (0.054) |
| Lower 95\% limit (l.d.) | 0.658 | 0.658 |
| Total catch (SD) | 16787 (2470) | 16745 (2459) |
| RMS catch difference | 16.8 | 16.9 |
| CV catch | 0.320 | 0.323 |
| Trial no. 3, $P_{0}=0.3 \mathrm{~K}, \mathrm{MSY} \%=1 \%$ |  |  |
| Mean final depletion (SD) | 0.469(0.056) | $0.554(0.074)$ |
| Mean lowest depletion (SD) | 0.299(0.005) | 0.300(0.002) |
| Lower 95\% limit (1.d.) | 0.300 | 0.300 |
| Total catch (SD) | 2805 (650) | 1964 (958) |
| RMS catch difference | 4.2 |  |
| CV catch | 0.514 | 1.291 |
| Trial no. 4, $P_{0}=0.3 \mathrm{~K}, \mathrm{MSY} \%=4 \%$ |  |  |
| Mean final depletion (SD) | 0.831 (0.056) | 0.837 (0.061) |
| Mean lowest depletion (SD) | 0.300 (0.000) | 0.300(0.000) |
| Lower 95\% limit (1.d.) | 0.300 | 0.300 |
| Total catch (SD) | 11039 (2871) | 10000 (3082) |
| RMS catch difference | 12.3 | 17.8 |
| CV catch | 0.593 | 0.802 |

## SUMMARY OF CONCLUSIONS

- The proposed management procedure provides satisfactory results for first stage screening trials and is ready for second stage testing.
- Further tuning of the various control parameters of the proposed procedure might be desirable to see if improved performance (in terms of smoother changes in catch limits) can be obtained. In particular, it may be useful to increase the number of MSY\% values allowed when fitting the population model to the data. However, this should first await the evaluation of the procedure's performance under second stage screening trials.
- Reducing the estimate of K by one or two standard errors results in losses in total catch for which slight increases in levels of depletion scarcely compensate. The more conservative approaches to which such reductions correspond thus seem unnecessary. An appropriate choice of the prior distribution for MSY\% (alone) provides adequate safeguards against the risk of overexploitation (assuming, naturally, that data of the precision specified for the trials continues to be forthcoming).
- An approach combining the estimator underlying the proposed management procedure with the NMP catch control law is unable to perform as well as the proposed procedure, because the NMP control law leads to inadequate data contrast for precise MSY\% estimation in development scenarios.
- Increasing the 'protection' level of the proposed management procedure results in lower total catches but higher final population sizes. The choice of this control parameter depends on what trade-off between these two attributes is desired.


## ACKNOWLEDGEMENTS

One of us (A.E.P.) acknowledges financial support from the Benguela Ecology Programme, Foundation for Research Development, South Africa.

## REFERENCES

International Whaling Commission. 1988. Comprehensive Assessment Workshop on Management, Reykjavik, 23-25 March 1987. Rep. int. Whal. Commn 38: 163-70.

International Whaling Commission. 1989. Report of the Comprehensive Assessment Workshop on Management Procedures, Lowestoft, February 1989. (Published in this volume)
Magnusson, K.G. and Stefansson, G. 1989. A feedback strategy to regulate catches from a whale stock. (Published in this volume).
Punt, A.E. and Butterworth, D.S. 1988. Investigations of some aspects of a potential whale stock management procedure. Paper SC/40/O 26, presented to the IWC Scientific Committee, May 1988. (Unpublished).
Punt, A.E. and Butterworth, D.S. 1989a. A proposed whale stock management procedure. Paper SC/F89/M2, presented to the Comprehensive Assessment Workshop on Management Procedures, Lowestoft, February 1989. (Unpublished).
Punt, A.E. and Butterworth, D.S. 1989b. Results of first stage screening trials for an improved proposed whale stock management procedure. Paper SC/41/O 19 presented to the IWC Scientific Committee, May 1989. (Unpublished).

# A Simulation Study on Management of Whale Stocks Considering Feedback Systems 

Kazumi Sakuramoto and Syoiti Tanaka<br>Tokyo University of Fisheries, 4-5-7, Konan, Minato-ku, Tokyo 108, Japan

Accepted May 1988


#### Abstract

A management procedure is examined using Monte-Carlo simulation. The procedure does not require information about the type of reproduction curve, the maximum sustainable yield or the level at which that is achieved. The whaling is regulated through a catch quota, which is increased or decreased according to the present level of CPUE relative to the target level of CPUE (TL) and its rate of change.

Control variables $g$ and $h$ are defined for deciding the catch quota. The stability of the stock-harvesting system is investigated for the following cases: (1) when the fixed values of TL, $g$ and $h$ are given; (2) when the number of whales suddenly decreases for some reason; (3) when the control variable $g$ and target level TL are incorporated into the feedback system; and (4) when the biological stock boundary does not coincide with the management area artificially set. The results of the simulation study are as follows: (1) some pairs of $g$ and $h$ values give high stability to the system; (2) it is very important whether or not the index of abundance is proportional to the population size; (3) stability is high when $g$ is changed according to the population level; (4) for some $g$ and $h$ values, the desirable TL value was successfully searched for automatically.


## INTRODUCTION

The New Management Procedure (NMP) of the International Whaling Commission (IWC) was adopted by the IWC in 1976. In the NMP, a Pella-Tomlinson (1969) type of reproduction curve is assumed and further, the maximum sustainable yield (MSY) and the MSYL, the population level producing MSY, are also assumed. However, they have not been observed for true populations and it is very difficult to confirm whether or not these assumptions are valid (Holt, 1985). It has been pointed out that if the values of MSY and MSYL assumed are invalid, the NMP may fail to control populations in a reliable manner (de la Mare, 1986a,b). Furthermore, for Southern Hemisphere minke whale stocks, this procedure cannot easily be applied because of the possiblity that the stocks have increased over time; the catch limits for those stocks have in fact been determined on the basis of replacement yield.

Tanaka (1986) stressed that the management of whale stocks is possible if the level of the stocks is monitored, even if there is insufficient information available for the NMP, such as the type of reproduction curve, the values of MSY and MSYL etc.

In Tanaka's method (1980; 1984), the target level of a population, TL, is chosen arbitrarily and then the catch quota is controlled such that the level of the population converges to the target level; i.e. the catch quota is increased or decreased taking into account the difference between current and target population levels, and whether or not the population is increasing or decreasing. Here, parameters $h$ and $g$ denote the weight given to the former and latter factors, respectively, for setting the catch quota. Tanaka (1980) noted as follows:
(1) that stability is high when TL is higher than MSYL but decreases when TL is lower than MSYL;
(2) that a time lag before the implementation of regulations has a serious effect on the stability of the system;
(3) that a small value of $h$ seems to be preferable and a value of $g$ around 1.0 seems to provide a high stability the ratio of $g$ to $h$ is connected with stability.

The aim of this simulation study is to investigate whether or not management is successful when Tanaka's method is used, if estimation errors in the stock level occur.

Firstly, considering Antarctic minke whale populations, the stability of the management procedures is investigated with fixed values of TL, $g$ and $h$. Three types of reproduction curve are assumed.

Secondly, the stability of the system when the number of whales suddenly decreases is investigated.

Thirdly, the control variable g is incorporated into the feedback systems in an attempt to improve the stability of the management system, especially for the case where the CPUE (catch per unit effort) is proportional to the square root of population size. Further, the decision process for TL is included into the feedback system of the management procedure; it is very important which values of $\mathrm{TL}, \mathrm{g}$ and h are chosen when the above method is applied in practice.

The above simulations assumed that the biological stock boundary completely coincides with the management area artificially set. In practice, however, this is unlikely to be true. In particular the management areas set for the Antarctic minke whale populations are originally those for fin whale populations. It is therefore possible that the true biological stock boundaries do not coincide with the management areas. A preliminary examination of the effects of differences between management areas and stock boundaries is given for two simple cases.

## THE BASIC STOCK-HARVESTING SYSTEM

## Dynamic model

A dynamic model of whale stocks is assumed as follows:

$$
\begin{equation*}
\mathbf{P}_{\mathbf{t}+1}=\left(\mathrm{P}_{\mathbf{t}}-\mathrm{Y}_{\mathbf{t}}\right) \exp (-\mathrm{M})+\mathrm{r}_{\mathrm{t}}-{ }_{l} \mathrm{P}_{\mathrm{t}}-l \tag{1}
\end{equation*}
$$

where; $P_{t}$ is the relative number of whales at the beginning of year $t ; Y_{t}$ is the relative catch in year $t ; M$ is the natural mortality coefficient, $(M=0.086) ; r_{t-l}$ is the rate of recruitment in year $t-l$; and $l$ is the age at sexual maturity, ( $l=7$ ).

* Originally presented as papers SC/38/O 10 and SC/M87/MI to the IWC Scientific Committee.

In this study, two types of reproduction curve are assumed:

$$
\begin{gather*}
\mathrm{r}_{\mathrm{t}-l}=\mathrm{M}\left[1+\mathrm{A}\left\{1-\left(\mathrm{P}_{\mathrm{t}}-\imath\right)^{\mathrm{n}}\right\}\right]  \tag{2}\\
\mathrm{r}_{\mathrm{t}-l}=\mathrm{M}(1+\mathrm{A}) /\left\{1+\mathrm{A}\left(\mathrm{P}_{\mathrm{t}}-l\right)^{\mathrm{n}}\right\} \tag{3}
\end{gather*}
$$

Equation (2) shows a Pella-Tomlinson type curve (Holt, 1985). The parameters $A$ and $n$ are related to the values of MSY and the density dependent compensation, respectively. We assumed two pairs of A and n values dependent on assumptions concerning supercompensation: $\mathrm{n}=2.39$ and $\mathrm{A}=1 / \mathrm{n}$ corresponds to non-supercompensation with MSY $=0.6$ [hereafter called the PT(N) model]; and $\mathrm{n}=2.39$ and $\mathrm{A}>1 / \mathrm{n}$ allows for supercompensation - A was taken to be 0.91 , the case where the recruitment reaches a maximum $10 \%$ greater than in the unexploited stock at $80 \%$ of the initial stock size [hereafter, called the $\mathrm{PT}(\mathrm{S})$ model].

Equation (3) shows a Shepherd type curve (Shepherd, 1982). For this, only the non-supercompensation case with MSY $=0.6$ was considered. The parameter values are $\mathrm{n}=$ 3.29 and $A=1 /(n-1)$ [hereafter, called $S D(N)$ model. See Fig. 1].

## Rule for deciding the catch quota

For the first ten years after exploitation, a constant level of catch, CY, is harvested. For simplicity, $4 \%$ of the initial stock size was used in this simulation. Detailed discussion of this value is given below.

After that period, the catch quota is decided according to the following rule:

$$
\begin{equation*}
\mathrm{Y}_{\mathrm{t}+1}=\left(1+\mathrm{hL}_{\mathrm{t}-l}+\mathrm{gK}_{\mathrm{t}-l}\right) \mathrm{Y}_{\mathrm{t}} \tag{4}
\end{equation*}
$$

where g and h are control variables; $\mathrm{L}_{\mathrm{t}-l}$ and $\mathrm{K}_{\mathrm{t}-l}$ are calculated using the values of CPUE during 11 years from $\mathrm{t}-l-5$ to $\mathrm{t}-l+4$. That is,

$$
\begin{gather*}
\mathrm{L}_{\mathrm{t}-\iota}=\left[\mathrm{E}\left(\mathrm{X}_{\mathrm{t}}-\imath\right) / \mathrm{TL}\right]-1  \tag{5}\\
\mathrm{~K}_{\mathrm{t}-\iota}=\mathrm{b} / \mathrm{E}\left(\mathrm{X}_{\mathrm{t}}-\imath\right) \tag{6}
\end{gather*}
$$

where TL denotes the target level of CPUE arbitrarily set and $X_{t}$ and $E\left(X_{t}\right)$ denote the value of CPUE at year $t$ and the mean value, respectively. That is,

$$
\begin{equation*}
\mathrm{E}\left(\mathrm{X}_{\mathrm{t}}-l\right)={ }_{\mathrm{t}-l-5}^{\mathrm{t}-l+5} \mathrm{X}_{\mathrm{i}} / 10 \tag{7}
\end{equation*}
$$

and $b$ denotes the slope of the regression line fitted to the CPUE series from from $t-l-5$ to $t-l+4$.

## CONDITIONS OF SIMULATION TESTS

## Catch per unit of effort

Two cases are assumed for the relationship between CPUE and population size: CPUE is proportional to the population size

$$
\begin{equation*}
X_{t}=k_{1} P_{t} \tag{8}
\end{equation*}
$$



Fig. 1. The three types of reproduction curve used for the simulation studies.
and CPUE is proportional to the square root of the population size

$$
\begin{equation*}
\mathrm{X}_{\mathrm{t}}=\mathrm{k}_{2} \sqrt{\mathrm{P}_{\mathrm{t}}} \tag{9}
\end{equation*}
$$

From equations (5) and (6), the $\mathrm{k}_{1}$ or $\mathrm{k}_{2}$ value is cancelled between the numerator and denominator; therefore the values $\mathrm{L}_{\mathrm{t}-l}$ and $\mathrm{K}_{\mathrm{t}-l}$ can be calculated regardless of the values of $\mathrm{k}_{1}$ or $\mathrm{k}_{2}$. It should be noted that the proportionality of CPUE to $P_{t}$ is always assumed to control the system even where the true relationship follows equation (9).

## Target level, TL

We investigated the case where the target level was set at 0.6 or 0.8 of the initial CPUE level. In some cases, 0.5 was used as the initial target level. The initial level of CPUE was set at unity.

## The control variables $g$ and $h$ set

Tanaka (1980) noted that a small value of $h$ seems to be preferable and a value of $g$ around 1.0 appears to give high stability ( $\mathrm{g}=1$ means that $10 \%$ of the catch quota is deducted when the CPUE decreases by $10 \%$ ). In this study, 1, 2, 3, 4 and 5 are used as $g$ values. Similarly, $h=0.1$ means that $2 \%$ of the catch quota is added when the CPUE is $20 \%$ higher than the target level. Values of $0.02,0.04$, $0.06,0.08$ and 0.10 are used as $h$ values.

## Noise

We assumed that the observed CPUE data contain noise uniformly distributed from -0.35 to +0.35 :

$$
\begin{equation*}
X_{t}(\text { obs })=X_{t}(\text { true })^{*}(1+\varepsilon),-0.35<\varepsilon<0.35 \tag{10}
\end{equation*}
$$

where $X_{t}(\mathrm{obs})$ and $\mathrm{X}_{\mathrm{t}}($ true $)$ are observed and true CPUE respectively. In some trials other levels of noise were also applied.

Twenty runs (occasionally 40) were conducted using different series of random numbers. Means of $P_{t}, X_{t}, Y_{t}$ and their standard deviations were calculated. The period of calculation is from 1 to 200 years.

Table 1
Conditions of simulations conducted

| Simulation | Table | Model | CPUE | Noise in | CPUE | TL set | Notes |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | 2 | PT(N) | $\propto{ }_{\text {H }}{ }_{\text {t }}$ | $-0.35<\varepsilon>0.35$ |  | 0.6 or 0.8 | Model |
|  | 3 | PT(S) |  |  |  |  |  |
|  | 4 | SD(N) |  |  |  | " |  |
|  | 5 | PT(N) | $\alpha /{ }^{\text {a }}$ | " |  | " | CPUE |
|  | 6 | PT(S) |  | " |  | " | $\alpha / \mathrm{P}_{\mathrm{t}}$ |
|  | 7 | SD(N) | " | " |  | " |  |
|  | 8 | PT(N) | $\underset{\alpha J \mathrm{P}_{\mathrm{t}}}{\mathrm{P}_{\mathrm{t}}}$ | $-0.050<\varepsilon$ | >0.50 | 0.6 | Noise |
| II | 9 | " | $\alpha P_{t}$ | $-0.35<\varepsilon>0.35$ |  |  | Perturbation |
| III | 10 | " | ${ }^{\prime \prime}$ | " |  | 0.5 or 0.8 | $\stackrel{\mathrm{g}}{\mathrm{TL}}$ |
|  | 11 12 | " | $\alpha \mathrm{P}_{\mathrm{t}}$ | " |  |  |  |
| 1V | 14 | " | ${ }_{\text {, }} \mathrm{P}_{\mathrm{t}}$ | " |  | 0.6 | $\begin{aligned} & 2 \text { stocks } \\ & \text { in } \end{aligned}$ |
|  | 15 | " |  |  |  | 0.8 |  |
|  | 16 |  | $\propto \sqrt{ } \mathrm{P}_{\mathrm{t}}$ |  |  | 0.6 | 1 Area |
|  | $17$ |  |  |  |  | 0.8 |  |
|  | 18 | " | ${ }_{\alpha} \mathrm{P}_{\mathrm{t}}$ | " |  | 0.6 or 0.8 | 1 stock |
|  | 19 | " | $\alpha / \mathrm{P}_{\mathrm{t}}$ |  |  |  | in 2 Areas |

## STABILITY OF STOCK-HARVESTING SYSTEM

## Indices of stability

In considering the stability of the stock-harvesting system, there are four points of interest.
(1) Mean of population size, PMEAN. This gives the average population level over 200 years:

$$
\begin{equation*}
\text { PMEAN }=\sum_{t=1}^{200} \overline{\mathrm{P}}_{\mathrm{t}} / 200 \tag{11}
\end{equation*}
$$

where, $\overline{\mathrm{P}}_{\mathrm{t}}$ is the average population size of the 20 runs for each year t .
(2) Minimum population level. This must be seriously considered in evaluating any management procedure because there is a possibility that the population may become very low or even extinct due to either a noisy or biased index of population size or a wrongly designed system. The minimum population size reached in the 20 runs is an important measure of the stability.
(3) Convergence of the population to the target level. The time taken for the trajectory to converge on the target level and the amplitude of the oscillations are another indication of stability. In this study, for simplicity, the stability of the system is judged by the magnitude of the area which is surrounded by curves of TL and $\bar{P}_{t}$, calculated using the formula

$$
\begin{equation*}
S=\sum_{t=1}^{200}\left|\bar{P}_{t}-\mathrm{TL}\right| \tag{12}
\end{equation*}
$$

(4) Catch level which can be continuously harvested. This is another index of the success of a procedure. The mean catch quota over the 200 years, YMEAN, is calculated as follows:

$$
\begin{equation*}
\text { YMEAN }=\sum_{t=1}^{200} \bar{Y}_{t} / 200 \tag{13}
\end{equation*}
$$

where $\bar{Y}_{t}$ denotes the average catch quota for the 20 runs in each year t .

## Table 2

Results for the case where the model is $\mathrm{PT}(\mathrm{N})$, CPUE is proportional to population and TL is set at 0.6 or 0.8

| $\mathrm{PT}(\mathrm{N}), \mathrm{CPUE} \propto \mathrm{P}_{\mathrm{t}}, \quad \mathrm{TL}=0.6$ |  |  |  |  |  | PT(N), CPUE $\propto$ P ${ }_{\text {t }}, \quad$ TL=0.8 |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathrm{g} / \mathrm{h}$ | 0.02 | 0.04 | 0.06 | 0.08 | 0.10 | g/h | 0.02 | 0.04 | 0.06 | 0.08 | 0.10 |
| PMEAN |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 0.324 | 0.508 | 0.580 | 0.607 | 0.611 | 1 | 0.587 | 0.701 | 0.743 | 0.763 | 0.772 |
| 2 | 0.576 | 0.593 | 0.604 | 0.606 | 0.601 | 2 | 0.686 | 0.732 | 0.755 | 0.768 | 0.775 |
| 3 | 0.640 | 0.623 | 0.617 | 0.614 | 0.612 | 3 | 0.730 | 0.754 | 0.768 | 0.777 | 0.782 |
| 4 | 0.685 | 0.653 | 0.637 | 0.629 | 0.624 | 4 | 0.763 | 0.776 | 0.783 | 0.789 | 0.792 |
| 5 | 0.723 | 0.684 | 0.660 | 0.646 | 0.638 | 5 | 0.791 | 0.796 | 0.800 | 0.802 | 0.803 |
| Minimum population |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 0.000 | 0.197 | 0.219 | 0.220 | 0.214 | 1 | 0.352 | 0.467 | 0.500 | 0.513 | 0.523 |
| 2 | 0.392 | 0.424 | 0.401 | 0.378 | 0.358 | 2 | 0.491 | 0.561 | 0.574 | 0.580 | 0.582 |
| 3 | 0.481 | 0.491 | 0.493 | 0.463 | 0.437 | 3 | 0.547 | 0.594 | 0.605 | 0.614 | 0.620 |
| 4 | 0.519 | 0.520 | 0.515 | 0.507 | 0.481 | 4 | 0.574 | 0.612 | 0.621 | 0.629 | 0.635 |
| 5 | 0.537 | 0.533 | 0.517 | 0.499 | 0.490 | 5 | 0.585 | 0.619 | 0.634 | 0.641 | 0.643 |
| $S=\Sigma\|\bar{P}(t)-T L\|$ |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 62.38 | 39.62 | 40.41 | 40.26 | 38.78 | 1 | 43.85 | 27.72 | 23.15 | 20.43 | 18.24 |
| 2 | 13.01 | 14.18 | 15.69 | 16.57 | 18.43 | 2 | 23.96 | 15.53 | 12.99 | 11.75 | 10.86 |
| 3 | 7.96 | 6.10 | 7.02 | 7.73 | 8.64 | 3 | 15.24 | 10.55 | 8.90 | 7.92 | 7.44 |
| 4 | 17.08 | 10.68 | 7.45 | 5.83 | 5.73 | 4 | 8.78 | 7.40 | 6.68 | 6.17 | 5.87 |
| 5 | 24.63 | 16.76 | 12.03 | 9.25 | 7.53 | 5 | 6.13 | 6.01 | 5.78 | 5.66 | 5.53 |
| YMEAN $\times 100$ |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 1.68 | 1.57 | 1.57 | 1.66 | 1.78 | 1 | 1.84 | 1.73 | 1.68 | 1.66 | 1.67 |
| 2 | 2.07 | 2.04 | 2.06 | 2.10 | 2.11 | 2 | 1.94 | 1.83 | 1.77 | 1.75 | 1.73 |
| 3 | 2.10 | 2.11 | 2.11 | 2.11 | 2.11 | 3 | 1.89 | 1.81 | 1.77 | 1.75 | 1.73 |
| 4 | 2.04 | 2.10 | 2.11 | 2.11 | 2.11 | 4 | 1.79 | 1.75 | 1.73 | 1.71 | 1.71 |
| 5 | 1.96 | 2.06 | 2.09 | 2.10 | 2.10 | 5 | 1.69 | 1.67 | 1.66 | 1.66 | 1.66 |

## SIMULATION I

## Sensitivity tests on TL, g, $h$ and reproduction curve

Sensitivity tests under the above conditions were carried out on TL, $g$ and $h$, the type of reproduction curve, the noise level and the CPUE/population size relationship.

Table 1 summarises the conditions for each run and the results are shown in Tables $2-8$.

Tables $2-4$ show the case where CPUE is proportional to population size. The reproduction curves are $\mathrm{PT}(\mathrm{N})$, $\mathrm{PT}(\mathrm{S})$ and $\mathrm{SD}(\mathrm{N})$ respectively and the noise level is set at 0.35 . Comparison among the tables shows that in the case

## Table 3

Results for the case where the model is $\mathrm{PT}(\mathrm{S})$, CPUE is proportional to population and TL is set at 0.6 or 0.8

| PT(S), CPUE^P ${ }_{t}$, TL=0.6 |  |  |  |  |  | PT(S), CPUE $\propto \mathrm{P}_{\mathrm{t}}, \quad$ TL=0.8 |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| g/h | 0.02 | 0.04 | 0.06 | 0.08 | 0.10 | g/h | 0.02 | 0.04 | 0.06 | 0.08 | 0.10 |
| PMEAN |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 0.641 | 0.627 | 0.617 | 0.604 | 0.600 | 1 | 0.779 | 0.787 | 0.791 | 0.793 | 0.795 |
| 2 | 0.694 | 0.649 | 0.634 | 0.626 | 0.621 | 2 | 0.801 | 0.802 | 0.802 | 0.801 | 0.801 |
| 3 | 0.733 | 0.679 | 0.654 | 0.641 | 0.633 | 3 | 0.821 | 0.817 | 0.814 | 0.812 | 0.810 |
| 4 | 0.767 | 0.709 | 0.677 | 0.658 | 0.647 | 4 | 0.841 | 0.834 | 0.828 | 0.824 | 0.821 |
| 5 | 0.798 | 0.741 | 0.703 | 0.680 | 0.665 | 5 | 0.860 | 0.852 | 0.845 | 0.839 | 0.834 |
| Minumum population |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 0.509 | 0.495 | 0.420 | 0.354 | 0.294 | 1 | 0.713 | 0.714 | 0.715 | 0.715 | 0.700 |
| 2 | 0.543 | 0.495 | 0.492 | 0.480 | 0.435 | 2 | 0.688 | 0.711 | 0.709 | 0.704 | 0.694 |
| 3 | 0.562 | 0.510 | 0.477 | 0.463 | 0.451 | 3 | 0.667 | 0.686 | 0.694 | 0.687 | 0.677 |
| 4 | 0.589 | 0.520 | 0.479 | 0.454 | 0.440 | 4 | 0.655 | 0.670 | 0.670 | 0.665 | 0.656 |
| 5 | 0.589 | 0.523 | 0.472 | 0.444 | 0.429 | 5 | 0.653 | 0.647 | 0.641 | 0.635 | 0.629 |
| $S=\Sigma\|\stackrel{\rightharpoonup}{P}(t)-T L\|$ |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 11.47 | 11.88 | 14.26 | 17.85 | 21.03 | 1 | 5.39 | 3.80 | 3.15 | 2.83 | 2.80 |
| 2 | 18.78 | 10.46 | 8.44 | 7.98 | 8.76 | 2 | 2.03 | 1.93 | 1.87 | 1.85 | 1.86 |
| 3 | 26.69 | 15.75 | 10.79 | 8.58 | 7.36 | 3 | 4.67 | 3.85 | 3.22 | 2.77 | 2.46 |
| 4 | 33.44 | 21.85 | 15.32 | 11.69 | 9.68 | 4 | 8.45 | 7.10 | 5.99 | 5.12 | 4.45 |
| 5 | 39.65 | 28.13 | 20.59 | 15.91 | 12.93 | 5 | 12.26 | 10.61 | 9.18 | 8.00 | 7.04 |
| YMEAN $\times 100$ |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 3.91 | 3.89 | 3.93 | 3.84 | 3.65 | 1 | 3.33 | 3.27 | 3.24 | 3.22 | 3.21 |
| 2 | 3.80 | 3.89 | 3.91 | 3.92 | 3.90 | 2 | 3.15 | 3.15 | 3.15 | 3.15 | 3.15 |
| 3 | 3.62 | 3.83 | 3.88 | 3.90 | 3.91 | 3 | 2.96 | 3.00 | 3.04 | 3.06 | 3.07 |
| 4 | 3.40 | 3.71 | 3.82 | 3.86 | 3.87 | 4 | 2.74 | 2.82 | 2.88 | 2.93 | 2.96 |
| 5 | 3.14 | 3.54 | 3.71 | 3.79 | 3.82 | 5 | 2.50 | 2.61 | 2.69 | 2.76 | 2.81 |

## Table 4

Results in the case where model is $\operatorname{SD}(\mathrm{N})$, CPUE is porportional to population and TL is set at 0.6 or 0.8

| $\mathrm{SD}(\mathrm{N}), \mathrm{CPUE} \propto \mathrm{P}_{\mathrm{t}}, \quad \mathrm{TL}=0.6$ |  |  |  |  |  | $\mathrm{SD}(\mathrm{N}), \mathrm{CPUE} \propto \mathrm{P}_{\mathrm{t}}, \mathrm{TL}=0.8$ |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathrm{g} / \mathrm{h}$ | 0.02 | 0.04 | 0.06 | 0.08 | 0.10 | $\mathrm{g} / \mathrm{h}$ | 0.02 | 0.04 | 0.06 | 0.08 | 0.10 |
| PMEAN |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 0.389 | 0.544 | 0.596 | 0.611 | 0.607 | 1 | 0.632 | 0.717 | 0.750 | 0.766 | 0.773 |
| 2 | 0.596 | 0.602 | 0.608 | 0.606 | 0.601 | 2 | 0.704 | 0.741 | 0.760 | 0.771 | 0.778 |
| 3 | 0.653 | 0.630 | 0.622 | 0.618 | 0.614 | 3 | 0.742 | 0.762 | 0.773 | 0.781 | 0.785 |
| 4 | 0.696 | 0.661 | 0.642 | 0.633 | 0.627 | 4 | 0.772 | 0.782 | 0.788 | 0.792 | 0.794 |
| 5 | 0.732 | 0.691 | 0.665 | 0.650 | 0.641 | 5 | 0.798 | 0.802 | 0.804 | 0.805 | 0.806 |
| Minimum population |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 0.035 | 0.231 | 0.241 | 0.237 | 0.230 | 1 | 0.414 | 0.503 | 0.529 | 0.539 | 0.546 |
| 2 | 0.430 | 0.450 | 0.422 | 0.394 | 0.371 | 2 | 0.525 | 0.583 | 0.593 | 0.598 | 0.601 |
| 3 | 0.501 | 0.507 | 0.507 | 0.476 | 0.447 | 3 | 0.569 | 0.609 | 0.620 | 0.627 | 0.633 |
| 4 | 0.531 | 0.529 | 0.510 | 0.501 | 0.488 | 4 | 0.587 | 0.625 | 0.634 | 0.641 | 0.644 |
| 5 | 0.544 | 0.536 | 0.514 | 0.493 | 0.482 | 5 | 0.593 | 0.624 | 0.644 | 0.646 | 0.648 |
| $S=\Sigma\|\bar{P}(t)-T L\|$ |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 49.47 | 36.23 | 37.06 | 35.99 | 35.68 | 1 | 35.04 | 22.59 | 18.84 | 16.59 | 14.93 |
| 2 | 9.60 | 11.51 | 12.99 | 14.32 | 16.30 | 2 | 20.36 | 13.13 | 10.68 | 9.57 | 8.89 |
| 3 | 10.67 | 6.16 | 6.21 | 6.71 | 7.46 | 3 | 12.76 | 9.00 | 7.64 | 6.70 | 6.22 |
| 4 | 19.25 | 12.16 | 8.47 | 6.56 | 5.68 | 4 | 7.05 | 6.35 | 5.79 | 5.37 | 5.12 |
| 5 | 26.47 | 18.16 | 13.08 | 10.05 | 8.18 | 5 | 5.60 | 5.48 | 5.38 | 5.27 | 5.10 |
| YMEAN $\times 100$ |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 1.85 | 1.77 | 1.81 | 1.93 | 2.03 | 1 | 2.03 | 1.90 | 1.84 | 1.82 | 1.82 |
| 2 | 2.26 | 2.24 | 2.26 | 2.28 | 2.27 | 2 | 2.08 | 1.97 | 1.91 | 1.88 | 1.86 |
| 3 | 2.26 | 2.28 | 2.28 | 2.29 | 2.29 | 3 | 2.00 | 1.93 | 1.89 | 1.86 | 1.85 |
| 4 | 2.19 | 2.26 | 2.28 | 2.28 | 2.28 | 4 | 1.89 | 1.85 | 1.83 | 1.82 | 1.81 |
| 5 | 2.08 | 2.20 | 2.25 | 2.27 | 2.27 | 5 | 1.77 | 1.76 | 1.75 | 1.75 | 1.76 |

of PT(S) (Table 3), PMEAN and YMEAN are high regardless of the $g$ and $h$ values. The minimum population size is high especially for the small $h$ values. The values of $S$ are small when g is 2 or 3 and h is $0.06-0.10$.

The values of PMEAN, YMEAN and minimum population for $\mathrm{PT}(\mathrm{N})$ (Table 2) are smaller than those of PT(S). Particulary when $g=1$, the minimum population size becomes seriously small. However, when $g=2$ or more or $\mathrm{TL}=0.8$, PMEAN and the minimum population become large. Comparing $\operatorname{SD}(\mathrm{N})$ (Table 4) with $\mathrm{PT}(\mathrm{N})$

## Table 5

Results for the case where the model is $\mathrm{PT}(\mathrm{N})$, CPUE is proportional to the square root of population and TL is set at 0.6 or 0.8

| $\operatorname{PT}(\mathrm{N}), \mathrm{CPUE} \times / \mathrm{P}_{\mathrm{t}}, \quad \mathrm{TL}=0.6$ |  |  |  |  |  | PT(N), CPUE $\propto / \mathrm{P}_{\mathrm{t}}, \quad$ TL=0.8 |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| g/h | h 0.02 | 0.04 | 0.06 | 0.08 | 0.10 | g/h | 0.02 | 0.04 | 0.06 | 0.08 | 0.10 |
| PMEAN |  |  |  |  |  |  |  |  |  |  |  |
|  | 0.154 | 0.134 | 0.121 | 0.111 | 0.104 | 1 | 0.196 | 0.246 | 0.420 | 0.522 | 0.577 |
| 2 | 0.214 | 0.170 | 0.145 | 0.129 | 0.118 | 2 | 0.441 | 0.526 | 0.577 | 0.608 | 0.627 |
| 3 | 0.379 | 0.318 | 0.315 | 0.325 | 0.329 | 3 | 0.580 | 0.603 | 0.619 | 0.630 | 0.638 |
|  | 0.497 | 0.418 | 0.397 | 0.402 | 0.414 | 4 | 0.651 | 0.652 | 0.652 | 0.653 | 0.653 |
| 5 | 0.575 | 0.483 | 0.440 | 0.424 | 0.420 | 5 | 0.703 | 0.693 | 0.685 | 0.679 | 0.674 |
| Minimum population |  |  |  |  |  |  |  |  |  |  |  |
|  | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 1 | 0.000 | 0.000 | 0.065 | 0.131 | 0.160 |
|  | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 2 | 0.000 | 0.256 | 0.290 | 0.293 | 0.294 |
|  | 0.000 | 0.053 | 0.071 | 0.032 | 0.000 | 3 | 0.256 | 0.344 | 0.388 | 0.383 | 0.374 |
|  | 0.193 | 0.187 | 0.190 | 0.157 | 0.107 | 4 | 0.345 | 0.384 | 0.419 | 0.425 | 0.413 |
| 5 | 0.298 | 0.269 | 0.232 | 0.231 | 0.174 | 5 | 0.379 | 0.410 | 0.439 | 0.437 | 0.426 |
| $S=\Sigma\|\bar{P}(t)-T L\|$ |  |  |  |  |  |  |  |  |  |  |  |
|  | 96.22 | 100.08 | 102.56 | 104.34 | 105.68 | 1 | 122.10 | 112.04 | 77.19 | 61.76 | 55.07 |
|  | 84.29 | 93.07 | 97.81 | 100.89 | 103.07 | 2 | 73.05 | 55.97 | 45.77 | 39.59 | 35.84 |
|  | 51.80 | 63.56 | 64.12 | 62.11 | 62.39 | 3 | 45.23 | 40.62 | 37.44 | 35.20 | 33.67 |
| 4 | 28.71 | 43.94 | 47.93 | 46.66 | 44.14 | 4 | 30.98 | 30.89 | 30.81 | 30.69 | 30.58 |
| 5 | 14.04 | 31.41 | 39.66 | 42.54 | 43.09 | 5 | 20.68 | 22.61 | 24.20 | 25.44 | 26.38 |
| YMEAN $\times 100$ |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 1.66 | 1.44 | 1.34 | 1.29 | 1.26 |  | 1.62 | 1.50 | 1.46 | 1.44 | 1.43 |
| 2 | 1.58 | 1.37 | 1.27 | 1.21 | 1.17 | 2 | 1.90 | 1.86 | 1.83 | 1.82 | 1.83 |
| 3 | 1.90 | 1.64 | 1.46 | 1.34 | 1.26 | 3 | 2.03 | 2.01 | 1.99 | 1.99 | 1.99 |
| 4 | 2.08 | 1.92 | 1.79 | 1.72 | 1.71 | 4 | 2.00 | 2.01 | 2.02 | 2.02 | 2.03 |
| 5 | 2.12 | 2.06 | 1.96 | 1.90 | 1.89 | 5 | 1.90 | 1.95 | 1.98 | 2.00 | 2.01 |

## Table 6

Results for the case where the model is $\mathrm{PT}(\mathrm{S})$, CPUE is proportional to the square root of population and TL is set at 0.6 or 0.8

| PT(S), CPUE $\propto / P_{t}, \quad$ TL=0.6 |  |  |  |  |  | PT(S), CPUE $\propto / \mathrm{P}_{\mathrm{t}}, \quad \mathrm{TL}=0.8$ |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| g/h | 0.02 | 0.04 | 0.06 | 0.08 | 0.10 | g/h | 0.02 | 0.04 | 0.06 | 0.08 | 0.10 |
| PMEAN |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 0.286 | 0.197 | 0.160 | 0.140 | 0.126 | 1 | 0.681 | 0.657 | 0.653 | 0.655 | 0.655 |
| 2 | 0.429 | 0.259 | 0.203 | 0.180 | 0.146 | 2 | 0.726 | 0.694 | 0.677 | 0.668 | 0.664 |
| 3 | 0.573 | 0.440 | 0.439 | 0.454 | 0.453 | 3 | 0.762 | 0.729 | 0.705 | 0.690 | 0.681 |
| 4 | 0.649 | 0.515 | 0.464 | 0.446 | 0.430 | 4 | 0.792 | 0.761 | 0.735 | 0.716 | 0.702 |
| 5 | 0.706 | 0.575 | 0.505 | 0.470 | 0.451 | 5 | 0.821 | 0.791 | 0.765 | 0.744 | 0.727 |
| Minimum population |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 1 | 0.534 | 0.485 | 0.486 | 0.457 | 0.414 |
| 2 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 2 | 0.562 | 0.520 | 0.471 | 0.451 | 0.450 |
| 3 | 0.240 | 0.089 | 0.124 | 0.105 | 0.030 | 3 | 0.527 | 0.529 | 0.491 | 0.453 | 0.431 |
| 4 | 0.331 | 0.205 | 0.160 | 0.181 | 0.153 | 4 | 0.498 | 0.502 | 0.498 | 0.455 | 0.424 |
| 5 | 0.379 | 0.263 | 0.181 | 0.165 | 0.168 | 5 | 0.477 | 0.479 | 0.479 | 0.447 | 0.412 |
| $S=\Sigma\|\bar{P}(t)-T L\|$ |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 74.87 | 90.56 | 96.90 | 100.49 | 102.88 | 1 | 25.12 | 29.88 | 30.54 | 30.20 | 30.23 |
| 2 | 48.93 | 79.20 | 89.07 | 92.90 | 99.22 | 2 | 15.95 | 22.35 | 25.87 | 27.54 | 28.44 |
| 3 | 24.38 | 44.57 | 42.62 | 42.19 | 46.83 | 3 | 8.92 | 15.52 | 20.20 | 23.18 | 25.04 |
| 4 | 16.10 | 31.28 | 38.66 | 40.91 | 43.22 | 4 | 3.15 | 9.12 | 14.28 | 18.07 | 20.74 |
| 5 | 21.12 | 22.33 | 31.70 | 36.68 | 39.62 | 5 | 5.04 | 3.19 | 8.22 | 12.47 | 15.75 |
| YMEAN $x 100$ |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 2.95 | 2.24 | 1.94 | 1.79 | 1.69 | 1 | 3.84 | 3.86 | 3.84 | 3.82 | 3.82 |
| 2 | 3.44 | 2.42 | 1.99 | 1.76 | 1.62 | 2 | 3.65 | 3.78 | 3.82 | 3.83 | 3.82 3.83 |
| 3 | 3.84 | 3.30 | 3.12 | 3.20 | 3.02 | 3 | 3.42 | 3.63 | 3.73 | 3.83 3.77 | 3.83 3.79 |
| 4 | 3.83 | 3.69 | 3.48 | 3.45 | 3.43 | 4 | 3.15 | 3.41 | 3.57 | 3.66 | 3.71 |
| 5 | 3.65 | 3.82 | 3.67 | 3.56 | 3.49 | 5 | 2.86 | 3.15 | 3.36 | 3.49 | 3.58 |

shows the former to give slightly larger PMEAN, minimum population and YMEAN values and usually slightly smaller S values. Contrary to Tanaka's (1980) expectation, $g$ values larger than 1.0 give high stability.

Tables $5-7$ show the results for the same runs as Tables 2-4 but with the CPUE proportional to the square root of the population size. In these cases, the probability of extinction is seriously high if $g$ is small. Therefore, large $g$ values and a high TL compared to MSYL should be chosen.

## Table 7

Results for the case where the model is $\mathrm{SD}(\mathrm{N}), \mathrm{CPUE}$ is proportional to the square root of population and TL is set at 0.6 and 0.8

| SD(N), CPUE $\alpha / P_{t}, \quad T L=0.6$ |  |  |  |  |  | $\mathrm{SD}(\mathrm{N}), \mathrm{CPUE} \alpha / \mathrm{P}_{\mathrm{t}}, \quad \mathrm{TL}=0.8$ |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.02 | 0.04 | 0.06 | 0.08 | 0.10 | g/h | 0.02 | 0.04 | 0.06 | 0.08 | 0.10 |
| PMEAN |  |  |  |  |  |  |  |  |  |  |  |
|  | 0.161 | 0.138 | 0.123 | 0.113 | 0.106 | 1 | 0.215 | 0.303 | 0.473 | 0.559 | 0.603 |
| 2 | 0.229 | 0.176 | 0.149 | 0.132 | 0.120 | 2 | 0.489 | 0.558 | 0.597 | 0.621 | 0.634 |
|  | 0.399 | 0.327 | 0.322 | 0.334 | 0.339 | 3 | 0.610 | 0.623 | 0.632 | 0.639 | 0.643 |
|  | 0.516 | 0.428 | 0.403 | 0.409 | 0.421 | 4 | 0.672 | 0.667 | 0.664 | 0.661 | 0.659 |
| 5 | 0.592 | 0.493 | 0.446 | 0.429 | 0.424 | 5 | 0.719 | 0.706 | 0.696 | 0.687 | 0.681 |
| Minimum population |  |  |  |  |  |  |  |  |  |  |  |
|  | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 1 | 0.000 | 0.000 | 0.100 | 0.158 | 0.181 |
|  | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 2 | 0.000 | 0.297 | 0.319 | 0.318 | 0.316 |
|  | 0.000 | 0.057 | 0.077 | 0.035 | 0.000 | 3 | 0.292 | 0.371 | 0.414 | 0.406 | 0.395 |
|  | 0.210 | 0.195 | 0.193 | 0.164 | 0.110 | 4 | 0.367 | 0.405 | 0.438 | 0.441 | 0.428 |
| 5 | 0.312 | 0.268 | 0.224 | 0.233 | 0.178 | 5 | 0.395 | 0.425 | 0.452 | 0.447 | 0.435 |
| $S=\Sigma\|\vec{P}(t)-T L\|$ |  |  |  |  |  |  |  |  |  |  |  |
|  | 94.91 | 99.33 | 102.06 | 103.97 | 105.40 | 1 | 118.24 | 100.53 | 67.50 | 55.89 | 49.92 |
|  | 81.55 | 91.84 | 97.09 | 100.40 | 102.71 | 2 | 63.48 | 49.55 | 41.72 | 37.12 | 34.37 |
|  | 47.90 | 61.93 | 62.70 | 60.77 | 61.41 | 3 | 39.31 | 36.62 | 34.80 | 33.50 | 32.64 |
|  | 25.30 | 42.30 | 46.80 | 45.46 | 42.93 | 4 | 26.74 | 27.75 | 28.52 | 29.04 | 29.39 |
|  | 11.68 | 29.73 | 38.64 | 41.77 | 42.51 | 5 | 17.38 | 19.96 | 22.11 | 23.79 | 25.08 |
| YMEAN $\times 100$ |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 1.73 | 1.49 | 1.38 | 1.33 | 1.29 | 1 | 1.72 | 1.63 | 1.60 | 1.60 | 1.62 |
| 2 | 1.69 | 1.44 | 1.32 | 1.24 | 1.20 | 2 | 2.09 | 2.06 | 2.05 | 2.04 | 2.06 |
| 3 | 2.06 | 1.75 | 1.54 | 1.40 | 1.32 | 3 | 2.20 | 2.19 | 2.18 | 2.18 | 2.19 |
| 4 | 2.25 | 2.06 | 1.91 | 1.85 | 1.85 | 4 | 2.14 | 2.17 | 2.19 | 2.19 | 2.20 |
| 5 | 2.28 | 2.22 | 2.10 | 2.03 | 2.02 | 5 | 2.02 | 2.08 | 2.13 | 2.15 | 2.17 |

## Table 8

Results for the case where the noise level is set at 0.5 , the model is $\mathrm{PT}(\mathbf{N})$ and CPUE is both proportional to population and the square root of population

| $\mathrm{PT}(\mathrm{N}), \mathrm{CPUE} \mathrm{P}_{\mathbf{t}}, \quad \mathrm{TL}=0.6$ |  |  |  |  |  | PT(N), CPUE $\alpha / \mathrm{P}_{\mathrm{t}}, \quad$ TL=0.6 |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| g/h | h 0.02 | 0.04 | 0.06 | 0.08 | 0.10 | g/h | 0.02 | 0.04 | 0.06 | 0.08 | 0.10 |
| PMEAN |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 0.325 | 0.504 | 0.576 | 0.602 | 0.606 | 1 | 0.154 | 0.134 | 0.121 | 0.112 | 0.105 |
| 2 | 0.580 | 0.595 | 0.606 | 0.607 | 0.603 | 2 | 0.223 | 0.173 | 0.147 | 0.131 | 0.122 |
| 3 | 0.651 | 0.630 | 0.622 | 0.618 | 0.615 | 3 | 0.405 | 0.333 | 0.323 | 0.327 | 0.327 |
| 4 | 0.706 | 0.669 | 0.649 | 0.637 | 0.630 | 4 | 0.538 | 0.447 | 0.415 | 0.412 | 0.414 |
| 5 | 0.754 | 0.709 | 0.680 | 0.662 | 0.650 | 5 | 0.633 | 0.531 | 0.473 | 0.446 | 0.435 |
| Minimum population |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 0.000 | 0.158 | 0.183 | 0.185 | 0.181 | 1 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 2 | 0.334 | 0.394 | 0.368 | 0.341 | 0.318 | 2 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 3 | 0.425 | 0.455 | 0.451 | 0.419 | 0.389 | 3 | 0.000 | 0.015 | 0.038 | 0.000 | 0.000 |
| 4 | 0.461 | 0.480 | 0.477 | 0.448 | 0.420 | 4 | 0.128 | 0.156 | 0.164 | 0.097 | 0.038 |
| 5 | 0.478 | 0.489 | 0.474 | 0.451 | 0.425 | 5 | 0.227 | 0.223 | 0.188 | 0.123 | 0.062 |
| $S=\Sigma\|\bar{P}(t)-T L\|$ |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 62.16 | 39.46 | 40.08 | 39.81 | 38.16 | 1 | 96.11 | 99.99 | 102.45 | 104.20 | 105.53 |
| 2 | 12.59 | 14.26 | 15.88 | 16.70 | 18.42 | 2 | 82.55 | 92.50 | 97.42 | 100.50 | 102.27 |
| 3 | 10.28 | 6.44 | 7.22 | 8.02 | 8.88 | 3 | 46.57 | 60.64 | 62.50 | 61.52 | 61.76 |
| 4 | 21.18 | 13.81 | 9.71 | 7.46 | 6.58 | 4 | 20.49 | 38.16 | 44.30 | 44.66 | 44.15 |
| 5 | 30.75 | 21.87 | 15.99 | 12.31 | 9.98 | 5 | 6.73 | 21.90 | 33.13 | 38.14 | 40.20 |
| YMEAN $\times 100$ |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 1.69 | 1.57 | 1.57 | 1.65 | 1.74 | 1 | 1.66 | 1.44 | 1.35 | 1.30 | 1.27 |
| 2 | 2.05 | 2.03 | 2.05 | 2.08 | 2.10 | 2 | 1.60 | 1.38 | 1.27 | 1.21 | 1.17 |
| 3 | 2.07 | 2.09 | 2.09 | 2.10 | 2.10 | 3 | 1.92 | 1.67 | 1.49 | 1.36 | 1.29 |
| 4 | 1.97 | 2.06 | 2.08 | 2.09 | 2.09 | 4 | 2.05 | 1.96 | 1.83 | 1.74 | 1.72 |
| 5 | 1.81 | 1.98 | 2.04 | 2.07 | 2.08 | 5 | 1.98 | 2.07 | 2.01 | 1.94 | 1.89 |

Table 8 shows the results of runs when the noise level is set at 0.5 . A comparison of Tables 2 and 8 shows the PMEAN values to be almost the same. If $g$ is large, PMEAN for noise level 0.5 is larger. However, for noise level 0.5 the minimum population size becomes smaller for all $g$ and $h$ values shown. The value of $S$ is slightly larger and YMEAN slightly smaller for the greater noise level. A comparison of Tables 8 and 5 reveals similar trends except that the $S$ value in Table 8 is slightly smaller. Some typical runs are illustrated in Fig. 2.

The above results suggest that in the case of $\mathrm{PT}(\mathrm{N})$, more caution is needed to control the population without risk of extinction. For that reason, the remaining simulations are confined to that case.

## SIMULATION II

## Perturbation of population level

This section considers the case where the population size decreases suddenly for some reason (hereafter, called perturbation). It is assumed that the number of sudden changes of population during the 200 years follows a Poisson distribution with a mean of 3 . Thus, the probability of the number of perturbations $x$ is given by,

$$
\begin{equation*}
P(x)=3 x e^{-3 / x!}, x=0,1, \ldots, 6 \tag{14}
\end{equation*}
$$

Values of $x$ larger than 6 are ignored. The time when the perturbation occurs is assumed to be random. We consider the situation that large perturbations seldom occur, but


Fig. 2. Some typical examples of trajectories of population, catch and CPUE.
that minor disturbances might occur frequently, and simply assume that the degree of perturbation, $y(\%)$, is linearly related to its number of occurrence:

$$
\begin{array}{ll}
\text { Assumption } 1 & y=2(11-x), x=1,2, \ldots, 6 \\
\text { Assumption } 2 & y=2(16-x), x=1,2, \ldots, 6 \tag{15}
\end{array}
$$

where $y(\%)$ denotes the maximum degree of perturbation. The actual percentage decrease varies randomly from 0 to $-y \%$.

Forty runs were carried out using different series of random numbers. The results are shown in Table 9. When the perturbations occur, PMEAN, the minimum population size and YMEAN become smaller. The $S$ value for small g values becomes larger and for large g values becomes smaller compared to the case without perturbation (see Table 2). The effect is particularly large for the minimum population size. The degree of the effect, of course, depends on how large and how often perturbations occur. However, if the $g$ value is set to be large, the effects of perturbation as assumed here seem not to be serious.

## SIMULATION III

## Incorporation of $\mathbf{g}$ into feedback system

The above results show that the management procedure sometimes fails if the control variable g is small and if the CPUE is proportional to the square root of the population size. In particular, for the case where TL is 0.6 , the possibility of extinction is high. The system must thus be modified to ensure that the possibility of extinction is small even if the CPUE is proportional to the square root of the population size. This is attempted here by incorporating $g$ into a feedback system, as described below.

Every ten years, an Auto-Regressive model is estimated for all the CPUE data accumulated up to that time:

$$
\begin{equation*}
\hat{X}_{t+1 \mid t}=\sum_{k=1}^{m} a_{k} X_{t-k+1} \tag{16}
\end{equation*}
$$

Table 9
Results for the case where perturbation occurs. TL is set at 0.6

| $\begin{gathered} \text { PT(N), CPUE } \propto P_{t}, \quad \text { TL }=0.6 \\ y=22-2 x \end{gathered}$ |  |  |  |  |  | PT(N), CPUE $\propto P_{t}, \quad T L=0.6$$y=32-2 x$ |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| g/h | 0.02 | 0.04 | 0.06 | 0.08 | 0.10 | $\mathrm{g} / \mathrm{h}$ | 0.02 | 0.04 | 0.06 | 0.08 | 0.10 |
| PMEAN |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 0.293 | 0.487 | 0.568 | 0.599 | 0.608 | 1 | 0.279 | 0.473 | 0.558 | 0.603 | 0.608 |
| 2 | 0.564 | 0.588 | 0.602 | 0.605 | 0.601 | 2 | 0.558 | 0.585 | 0.600 | 0.604 | 0.601 |
| 3 | 0.629 | 0.617 | 0.614 | 0.612 | 0.610 | 3 | 0.625 | 0.615 | 0.613 | 0.611 | 0.609 |
| 4 | 0.674 | 0.647 | 0.633 | 0.626 | 0.622 | 4 | 0.671 | 0.645 | 0.632 | 0.625 | 0.621 |
| 5 | 0.711 | 0.676 | 0.655 | 0.643 | 0.635 | 5 | 0.708 | 0.674 | 0.654 | 0.642 | 0.634 |
| Minimum population |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 0.000 | 0.134 | 0.162 | 0.167 | 0.165 | 1 | 0.000 | 0.092 | 0.131 | 0.139 | 0.141 |
| 2 | 0.349 | 0.341 | 0.340 | 0.325 | 0.309 | 2 | 0.300 | 0.295 | 0.295 | 0.288 | 0.269 |
| 3 | 0.438 | 0.428 | 0.419 | 0.398 | 0.378 | 3 | 0.384 | 0.377 | 0.365 | 0.346 | 0.328 |
| 4 | 0.485 | 0.472 | 0.463 | 0.444 | 0.425 | 4 | 0.429 | 0.420 | 0.405 | 0.387 | 0.371 |
| 5 | 0.508 | 0.493 | 0.484 | 0.478 | 0.462 | 5 | 0.453 | 0.441 | 0.435 | 0.419 | 0.404 |
| $S=\Sigma\|\bar{P}(t)-T L\|$ |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 68.08 | 40.64 | 40.71 | 40.66 | 39.27 | 1 | 70.70 | 41.03 | 40.47 | 40.40 | 39.19 |
| 2 | 14.76 | 15.39 | 16.82 | 17.30 | 18.92 | 2 | 15.73 | 15.98 | 17.37 | 17.70 | 18.97 |
| 3 | 5.90 | 6.25 | 7.43 | 8.15 | 8.98 | 3 | 5.55 | 6.48 | 7.71 | 8.43 | 9.24 |
| 4 | 14.88 | 9.42 | 6.66 | 5.48 | 5.60 | 4 | 14.11 | 9.00 | 6.39 | 5.46 | 5.63 |
| 5 | 22.25 | 15.25 | 11.04 | 8.54 | 7.00 | 5 | 21.54 | 14.86 | 10.79 | 8.36 | 6.85 |
| YMEAN $\times 100$ |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 1.61 | 1.50 | 1.49 | 1.56 | 1.61 | 1 | 1.58 | 1.46 | 1.44 | 1.50 | 1.58 |
| 2 | 1.98 | 1.95 | 1.97 | 2.01 | 2.03 | 2 | 1.93 | 1.90 | 1.91 | 1.95 | 1.97 |
| 3 | 2.03 | 2.03 | 2.03 | 2.04 | 2.04 | 3 | 1.98 | 1.98 | 1.98 | 1.98 | 1.99 |
| 4 | 1.98 | 2.02 | 2.03 | 2.04 | 2.04 | 4 | 1.93 | 1.97 | 1.98 | 1.99 | 1.99 |
| 5 | 1.90 | 1.99 | 2.02 | 2.03 | 2.03 | 5 | 1.85 | 1.93 | 1.97 | 1.98 | 1.98 |

The degree of $m$ is decided by AIC (Akaike information criterion, Akaike, 1971). Using the estimated values of $m$ and $a_{1,2}, \ldots a_{m}$, the control variable $g$ is modified according to the following rule:

$$
\begin{align*}
& E\left(X_{t}\right)=\sum_{i=t-9}^{t} X_{i} / 10 \\
& E\left(\hat{X}_{t}\right)=\sum_{i=t+1}^{t+10} X_{i \mid t} / 10 \tag{17}
\end{align*}
$$

where $E\left(X_{t}\right)$ and $E\left(\hat{X}_{t}\right)$ are the mean values of observed and predicted CPUE over the 10 years respectively. The valueg is modified only when the CPUE level forecasted is less than TL (Fig. 3). That is

$$
\begin{align*}
& 0.9 \mathrm{TL}<\mathrm{E}\left[\hat{\mathrm{X}}_{\mathrm{t}}\right]<\mathrm{TL} \text { then } \mathrm{g}^{\prime}=2 \mathrm{~g} \\
& \mathrm{E}\left[\hat{\mathrm{X}}_{\mathrm{t}}\right] \leq 0.9 \mathrm{TL} \text { then } \mathrm{g}^{\prime}=4 \mathrm{~g} \tag{18}
\end{align*}
$$



Fig. 3. Examples of past and projected trajectories.

Only an increase in the value of $g$ is considered here. Table 10 shows the results for $\mathrm{TL}=0.6$. A comparison with Table 2 shows an improvement at $g=1$. A comparison with Table 5 shows an improvement in the PMEAN and S values for all $g$ and $h$ values. However, improvement in the minimum population size occurs only for small $h$ values.

## Feedback system to renew the target level automatically

The results from the previous section show that stability of the system is heavily dependent on the CPUE level set as the target level. However, in practice, the population abundance is rarely known exactly. It is therefore desirable to modify the target level automatically in the feedback system without assuming any type of reproduction curve.

## Basic concept for modifying TL

As mentioned before, the catch quota, $\mathbf{Y}_{t}$, is decided by the level and the trend in CPUE. Therefore, if $\mathrm{Y}_{\mathrm{t}}$ is larger than the replacement yield, the population size in the next year, $P_{t+1}$, will be smaller than $P_{t}$, i.e. $X_{t+1}$ is smaller than $X_{t}$, and vice versa.

Table 10
Results for the case where $g$ is incorporated into the feedback system

| PT(N), CPUE $\alpha \mathrm{P}_{t}, \quad \mathrm{TL}=0.6$ |  |  |  |  |  | PT(N), CPUE $\propto / \mathrm{P}_{\mathrm{t}}, \quad \mathrm{TL}=0.6$ |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| g/h | 0.02 | 0.04 | 0.06 | 0.08 | 0.10 | g/h | 0.02 | 0.04 | 0.06 | 0.08 | 0.10 |
| PMEAN |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 0.506 | 0.560 | 0.588 | 0.593 | 0.589 | 1 | 0.265 | 0.244 | 0.290 | 0.309 | 0.261 |
| 2 | 0.595 | 0.597 | 0.603 | 0.601 | 0.600 | 2 | 0.475 | 0.449 | 0.426 | 0.424 | 0.417 |
| 3 | 0.649 | 0.629 | 0.627 | 0.622 | 0.618 | 3 | 0.577 | 0.530 | 0.523 | 0.491 | 0.486 |
| 4 | 0.686 | 0.658 | 0.643 | 0.634 | 0.630 | 4 | 0.633 | 0.617 | 0.583 | 0.563 | 0.548 |
| 5 | 0.725 | 0.686 | 0.665 | 0.651 | 0.643 | 5 | 0.660 | 0.652 | 0.645 | 0.606 | 0.606 |
| Minimum population |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 0.276 | 0.323 | 0.346 | 0.312 | 0.311 | 1 | 0.000 | 0.000 | 0.013 | 0.020 | 0.000 |
| 2 | 0.381 | 0.391 | 0.402 | 0.413 | 0.386 | 2 | 0.186 | 0.165 | 0.205 | 0.159 | 0.144 |
| 3 | 0.460 | 0.465 | 0.446 | 0.454 | 0.445 | 3 | 0.230 | 0.203 | 0.177 | 0.168 | 0.148 |
| 4 | 0.519 | 0.521 | 0.524 | 0.507 | 0.481 | 4 | 0.228 | 0.295 | 0.080 | 0.059 | 0.046 |
| 5 | 0.537 | 0.533 | 0.521 | 0.504 | 0.496 | 5 | 0.306 | 0.268 | 0.213 | 0.146 | 0.000 |
| $S=\Sigma\|\bar{P}(t)-T L\|$ |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 26.05 | 23.47 | 23.67 | 25.18 | 26.35 | 1 | 73.91 | 78.00 | 68.72 | 65.23 | 74.28 |
| 2 | 10.67 | 11.78 | 11.97 | 11.64 | 11.93 | 2 | 32.12 | 37.13 | 41.61 | 42.01 | 43.22 |
| 3 | 9.73 | 6.93 | 8.10 | 8.61 | 8.35 | 3 | 13.95 | 21.22 | 22.48 | 28.69 | 29.58 |
| 4 | 17.20 | 11.65 | 8.66 | 6.89 | 6.78 | 4 | 14.04 | 15.42 | 13.49 | 14.63 | 17.92 |
| 5 | 25.08 | 17.13 | 12.99 | 10.12 | 8.65 | 5 | 12.43 | 15.04 | 20.32 | 13.95 | 16.32 |
| YMEAN $x 100$ |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 1.94 | 1.92 | 1.97 | 2.04 | 2.06 | 1 | 1.55 | 1.36 | 1.33 | 1.24 | 1.15 |
| 2 | 2.07 | 2.06 | 2.08 | 2.08 | 2.08 | 2 | 1.97 | 1.92 | 1.89 | 1.90 | 1.88 |
| 3 | 2.09 | 2.09 | 2.09 | 2.10 | 2.09 | 3 | 2.01 | 1.98 | 1.98 | 1.97 | 1.96 |
| 4 | 2.04 | 2.09 | 2.10 | 2.09 | 2.09 | 4 | 1.97 | 1.95 | 1.96 | 1.98 | 1.93 |
| 5 | 1.95 | 2.05 | 2.08 | 2.09 | 2.09 | 5 | 1.94 | 1.87 | 1.85 | 1.89 | 1.86 |

The pairs of $\left(\mathrm{X}_{\mathrm{t}-l}, \mathrm{Y}_{\mathrm{t}}\right)$ are separated into two groups according to the following rule:
$X_{t+1}-X_{t}>0$ then $\left(X_{t-l}, Y_{t}\right) \varepsilon$ Group $U$
$X_{t+1}-X_{t}<0$ then $\left(X_{t-l}, Y_{t}\right) \varepsilon$ Group $V$
On average, the production curve can be considered to be located on the line which separates the domains of these two groups. Therefore, we calculate the discriminant function of these two groups over a certain period of years. The calculation was made every 10 years using the 40 pairs of data, $\left(\mathrm{X}_{\mathrm{t}-1}, \mathrm{Y}_{\mathrm{t}}\right)$, prior to year t (in the case where $\mathrm{t}-40<0$, the discriminant function was calculated with data less than 40).

## Rules for adjusting the target level

Alpha denotes the slope of the discriminant function and $\mathrm{E}\left(\mathrm{X}_{\mathrm{k}}\right)$ denotes the mean CPUE during the periods when the discriminant function has been calculated (Fig. 4)
(1) If $E\left(X_{k}\right)>T L$ and $\alpha \leq 0$, then the peak of the reproduction curve is considered to be at the left hand side of $E\left(X_{k}\right)$. The TL at the left hand side of $E\left(X_{k}\right)$ is already set, thus no renewal of TL is needed.
(2) If $\mathrm{E}\left(\mathrm{X}_{\mathrm{k}}\right)>\mathrm{TL}$ and $\alpha>0$, then the peak of the reproduction curve is considered to be at the right hand side of $E\left(X_{k}\right)$. Therefore, a larger TL should be set:

$$
\begin{equation*}
\mathrm{TL}:=\mathrm{TL}(1+\mathrm{w} \alpha) \tag{20}
\end{equation*}
$$

where, $w$ is the weight used to accelerate the speed of adjustment and it is set arbitrarily

$$
\begin{align*}
& w=5(T L-0.5)+1, \mathrm{TL} \geq 0.5 \\
& \mathrm{w}=-5(\mathrm{TL}-0.5)+1, \mathrm{TL}<0.5 \tag{21}
\end{align*}
$$

However, modification of TL is limited to within $10 \%$ of the present TL in one renewal.
(3) If $\mathrm{E}\left(\mathrm{X}_{\mathrm{k}}\right)<\mathrm{TL}$ and $\alpha \leq 0$, then the peak of the reproduction curve is considered to be at the left hand side of $E\left(X_{k}\right)$. Therefore, a smaller TL should be set following equation (20).
(4) If $\mathrm{E}\left(\mathrm{X}_{\mathrm{k}}\right)<\mathrm{TL}$ and $\alpha>0$, then the peak of the reproduction curve is considered to be at the right hand side of $E\left(X_{k}\right)$. TL has been set at the right hand side of $\mathrm{E}\left(\mathrm{X}_{\mathrm{k}}\right)$ and thus no modification is needed.


Fig. 4. Schematic illustration of the concept of Renewing TL. Solid lines show the discriminant function calculated and dotted curves show the actual reproduction curve.

Table 11
Some examples where the TL was successfully adjusted

|  | $\begin{gathered} g=3 \\ h=0.02 \\ \hline \end{gathered}$ | $\begin{gathered} g=4 \\ h=0.02 \end{gathered}$ | $\begin{gathered} g=4 \\ h=0.04 \\ \hline \end{gathered}$ | $\begin{gathered} g=5 \\ h=0.04 \end{gathered}$ | $\begin{gathered} g=5 \\ h=0.06 \end{gathered}$ |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Year | TL | TL | TL | TL | TL | SD |
| 11-20 | 0.5375 | 0.5277 | 0.5315 | 0.5310 | 0.5297 | 0.01394 |
| 21-30 | 0.5456 | 0.5511 | 0.5494 | 0.5565 | 0.5543 | 0.03007 |
| 31-40 | 0.5499 | 0.5663 | 0.5566 | 0.5736 | 0.5626 | 0.05215 |
| 41-50 | 0.5637 | 0.5958 | 0.5711 | 0.5972 | 0.5773 | 0.06730 |
| 51-60 | 0.5702 | 0.6150 | 0.5853 | 0.6071 | 0.5892 | 0.05834 |
| 61-70 | 0.5733 | 0.6360 | 0.5950 | 0.6173 | 0.6004 | 0.06252 |
| 71-80 | 0.5852 | 0.6465 | 0.6019 | 0.6306 | 0.6065 | 0.06468 |
| 81-90 | 0.5927 | 0.6564 | 0.6030 | 0.6375 | 0.6101 | 0.06076 |
| 91-100 | 0.5896 | 0.6583 | 0.5946 | 0.6275 | 0.5970 | 0.05998 |
| 101-110 | 0.5843 | 0.6591 | 0.6074 | 0.6365 | 0.6024 | 0.06284 |
| 111-120 | 0.5743 | 0.6558 | 0.6017 | 0.6359 | 0.6016 | 0.05894 |
| 121-130 | 0.5785 | 0.6571 | 0.6012 | 0.6471 | 0.6109 | 0.06323 |
| 131-140 | 0.5800 | 0.6645 | 0.6039 | 0.6519 | 0.6132 | 0.07171 |
| 141-150 | 0.5802 | 0.6568 | 0.6037 | 0.6474 | 0.6192 | 0.08437 |
| 151-160 | 0.5799 | 0.6602 | 0.6165 | 0.6569 | 0.6231 | 0.07504 |
| 161-170 | 0.5814 | 0.6596 | 0.6174 | 0.6560 | 0.6238 | 0.07926 |
| 171-180 | 0.5783 | 0.6479 | 0.6067 | 0.6514 | 0.6179 | 0.06844 |
| 181-190 | 0.5828 | 0.6463 | 0.6006 | 0.6543 | 0.6214 | 0.06697 |
| 191-200 | 0.5865 | 0.6559 | 0.6038 | 0.6553 | 0.6156 | 0.06328 |
|  | $\begin{gathered} g=1 \\ h=0.02 \end{gathered}$ | $\begin{gathered} g=1 \\ h=0.04 \end{gathered}$ | $\begin{gathered} g=2 \\ h=0.02 \end{gathered}$ | $\begin{gathered} g=2 \\ h=0.04 \end{gathered}$ |  |  |
| Year | TL | TL | TL | TL | TL | SD |
| 11-20 | 0.8000 | 0.8000 | 0.8000 | 0.8000 | 0.8715 | 0.02086 |
| 21-30 | 0.8000 | 0.8000 | 0.8000 | 0.8000 | 0.8330 | 0.03946 |
| 31-40 | 0.8000 | 0.8000 | 0.8000 | 0.8000 | 0.8174 | 0.04177 |
| 41-50 | 0.8000 | 0.8000 | 0.8000 | 0.8000 | 0.8115 | 0.04861 |
| 51-60 | 0.8000 | 0.8000 | 0.7853 | 0.7970 | 0.7922 | 0.06404 |
| 61-70 | 0.8000 | 0.8000 | 0.7586 | 0.7721 | 0.7681 | 0.07474 |
| 71-80 | 0.8000 | 0.7501 | 0.7201 | 0.7235 | 0.7396 | 0.07073 |
| 81-90 | 0.8000 | 0.6781 | 0.6914 | 0.7004 | 0.7338 | 0.06666 |
| 91-100 | 0.7857 | 0.6641 | 0.6671 | 0.6761 | 0.7192 | 0.06069 |
| 101-110 | 0.7200 | 0.6641 | 0.6604 | 0.6838 | 0.7158 | 0.05682 |
| 111-120 | 0.6638 | 0.6641 | 0.6621 | 0.6986 | 0.7201 | 0.04831 |
| 121-130 | 0.6462 | 0.6641 | 0.6680 | 0.7054 | 0.7315 | 0.05475 |
| 131-140 | 0.6462 | 0.6761 | 0.6630 | 0.7159 | 0.7274 | 0.04985 |
| 141-150 | 0.6462 | 0.6901 | 0.6583 | 0.7129 | 0.7299 | 0.05291 |
| 151-160 | 0.6462 | 0.6987 | 0.6734 | 0.7185 | 0.7313 | 0.06327 |
| 161-170 | 0.6462 | 0.7210 | 0.6708 | 0.7300 | 0.7329 | 0.06038 |
| 171-180 | 0.6462 | 0.7315 | 0.6736 | 0.7232 | 0.7258 | 0.06497 |
| 181-190 | 0.6462 | 0.7419 | 0.6719 | 0.7207 | 0.7324 | 0.05675 |
| 191-200 | 0.6637 | 0.7613 | 0.6759 | 0.7221 | 0.7331 | 0.06127 |



Fig. 5. Schematic illustration of the relationship between biological stock boundaries and management areas.

Table 12
Results for the case where TL is adjusted

| $\mathrm{PT}(\mathrm{N}), \mathrm{CPUE} \times_{\mathrm{P}}, \quad$ TL=0.5 |  |  |  |  |  | $\mathrm{PT}(\mathrm{N}), \mathrm{CPUE} \mathrm{\propto} \mathrm{P}_{t}, \quad \mathrm{TL}=0.8$ |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| g/h | 0.02 | 0.04 | 0.06 | 0.08 | 0.10 | g/h | 0.02 | 0.04 | 0.06 | 0.08 | 0.10 |
| PMEAN |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 0.191 | 0.162 | 0.168 | 0.173 | 0.163 | 1 | 0.579 | 0.672 | 0.705 | 0.707 | 0.720 |
| 2 | 0.479 | 0.486 | 0.498 | 0.497 | 0.488 | 2 | 0.663 | 0.704 | 0.721 | 0.719 | 0.717 |
| 3 | 0.605 | 0.548 | 0.524 | 0.514 | 0.508 | 3 | 0.716 | 0.735 | 0.742 | 0.747 | 0.752 |
| 4 | 0.679 | 0.626 | 0.585 | 0.549 | 0.538 | 4 | 0.759 | 0.771 | 0.777 | 0.780 | 0.779 |
| 5 | 0.720 | 0.676 | 0.638 | 0.601 | 0.580 | 5 | 0.788 | 0.794 | 0.794 | 0.795 | 0.799 |
| Minimum population |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 1 | 0.352 | 0.467 | 0.500 | 0.513 | 0.523 |
| 2 | 0.268 | 0.275 | 0.220 | 0.146 | 0.120 | 2 | 0.491 | 0.561 | 0.560 | 0.539 | 0.467 |
| 3 | 0.430 | 0.406 | 0.368 | 0.315 | 0.269 | 3 | 0.547 | 0.594 | 0.593 | 0.588 | 0.567 |
| 4 | 0.479 | 0.450 | 0.437 | 0.394 | 0.327 | 4 | 0.574 | 0.612 | 0.621 | 0.629 | 0.635 |
| 5 | 0.503 | 0.471 | 0.453 | 0.432 | 0.387 | 5 | 0.585 | 0.617 | 0.634 | 0.641 | 0.647 |
| $S=\Sigma\|\bar{P}(t)-T L\|$ |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 74.42 | 79.32 | 77.61 | 76.21 | 77.95 | 1 | 32.10 | 25.40 | 16.58 | 19.57 | 13.37 |
| 2 | 19.50 | 22.08 | 29.90 | 29.12 | 33.56 | 2 | 7.82 | 9.00 | 22.32 | 10.77 | 13.47 |
| 3 | 12.85 | 10.77 | 12.13 | 13.81 | 16.67 | 3 | 4.76 | 10.07 | 6.66 | 5.03 | 5.96 |
| 4 | 8.19 | 10.91 | 17.31 | 14.47 | 13.71 | 4 | 9.52 | 4.91 | 6.40 | 4.61 | 7.32 |
| 5 | 6.39 | 7.87 | 11.62 | 13.69 | 21.49 | 5 | 6.06 | 5.42 | 8.92 | 3.80 | 3.95 |
| YMEAN $\times 100$ |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 1.61 | 1.36 | 1.24 | 1.17 | 1.13 | 1 | 1.89 | 1.90 | 1.89 | 1.94 | 1.90 |
| 2 | 1.99 | 1.91 | 1.97 | 2.04 | 2.04 | 2 | 2.03 | 1.96 | 1.93 | 1.94 | 1.93 |
| 3 | 2.10 | 2.09 | 2.08 | 2.07 | 2.05 | 3 | 1.95 | 1.90 | 1.89 | 1.88 | 1.86 |
| 4 | 2.02 | 2.08 | 2.10 | 2.09 | 2.08 | 4 | 1.82 | 1.78 | 1.77 | 1.76 | 1.77 |
| 5 | 1.92 | 2.03 | 2.07 | 2.09 | 2.09 | 5 | 1.70 | 1.68 | 1.69 | 1.70 | 1.68 |

## Examples of the results of simulations

Table 11 shows some examples where TL has been successfully adjusted for cases where the initial TL is set at 0.5 or 0.8 . When the initial TL is set at 0.5 , a combination of large g and small h values seems to be preferable. However, if $g$ is set at 1 , the modified TL in fact becomes worse. When the initial TL is set at 0.8 , a combination of small g and small h values seems to be preferable.

The results for an initial TL of 0.5 suggest that the TL would be modified more quickly and closer to the true TL than when the initial TL is 0.8 . If the initial TL is set at 0.5 , the population level rapidly decreases around the true MSYL. Thus the slope of discriminant function is much steeper than for $\mathrm{TL}=0.8$. Similar reasoning explains why a small g value is suitable for $\mathrm{TL}=0.8$ as a small g value produces a lower population level.

Table 12 shows PMEAN, the minimum population size, $S$ and YMEAN, when TL is modified. In comparison with the results of Table 2, for $\mathrm{TL}=0.5$, the system seems to behave well except for the case of small $g$ values. For $\mathrm{TL}=0.8$, the results are similar to those of Table 2. However, YMEAN becomes larger by modifying TL, especially in the case of small $g$ values.

## SIMULATION IV

## Two stocks coexist in one management area

In this section, cases where the biological stock boundary does not coincide with the management area are considered. In particular, the case where two biologically separate stocks (Stocks A and B) coexist in the management area (Fig. 5). The stability of the system if A and $B$ are treated as a single stock under this management procedure is investigated, assuming a high fishing intensity on stock B in order to examine a 'serious' case. Exploitation in the simulation started from the initial population size and the reproductive system is the same for both stocks, i.e. a Pella-Tomlinson model with no supercompensation and an MSYL set at 0.6 or 0.8 of the initial CPUE.

Table 13
Relationship between the ratio of population size and the fishing intensity between assumed stocks $A$ and $B$

| Ratio of population size |  |  | Proportion of catches |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Stock $\mathbf{A}$ | $:$ | Stock $\mathbf{B}$ | Stock A | $:$ | Stock B |
| 1 | $\vdots$ | 1 | 0.30 | $:$ | 0.70 |
| 1 | $\vdots$ | $1 / 2$ | 0.42 | $:$ | 0.58 |
| 1 | $\vdots$ | $1 / 3$ | 0.51 | $:$ | 0.49 |
| 1 | $\vdots$ | $1 / 4$ | 0.60 | $:$ | 0.40 |
| 1 | $:$ | $1 / 9$ | 0.90 | $:$ | 0.10 |

Table 14
Results for the case where two stocks coexist in one management area. Stock B has suffered high fishing intensity. TL is set at 0.6. CPUE is proportional to population.

| PT(N), CPUE $\alpha \mathrm{P}_{\mathrm{t}}, \quad \mathrm{TL}=0.6$ |  |  |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Stock A |  |  |  |  | Stock B |  |  |  |  |  |
| g/h 0.02 | 0.04 | 0.06 | 0.08 | 0.10 | g/h | 0.02 | 0.04 | 0.06 | 0.08 | 0.10 |
| PMEAN |  |  |  |  |  |  |  |  |  |  |
| 10.820 | 0.857 | 0.872 | 0.876 | 0.873 | 1 | 0.291 | 0.377 | 0.442 | 0.477 | 0.488 |
| 20.866 | 0.870 | 0.872 | 0.872 | 0.871 | 2 | 0.462 | 0.471 | 0.480 | 0.486 | 0.488 |
| 30.884 | 0.879 | 0.877 | 0.875 | 0.872 | 3 | 0.552 | 0.530 | 0.518 | 0.512 | 0.508 |
| 40.899 | 0.889 | 0.884 | 0.881 | 0.879 | 4 | 0.617 | 0.581 | 0.558 | 0.543 | 0.534 |
| 50.913 | 0.901 | 0.893 | 0.888 | 0.885 | 5 | 0.670 | 0.627 | 0.598 | 0.578 | 0.563 |
| Minumum population |  |  |  |  |  |  |  |  |  |  |
| 10.664 | 0.651 | 0.642 | 0.630 | 0.614 | 1 | 0.137 | 0.164 | 0.176 | 0.181 | 0.181 |
| 20.744 | 0.730 | 0.716 | 0.703 | 0.691 | 2 | 0.298 | 0.321 | 0.326 | 0.314 | 0.300 |
| 30.773 | 0.760 | 0.747 | 0.735 | 0.722 | 3 | 0.371 | 0.382 | 0.371 | 0.363 | 0.359 |
| $4 \quad 0.788$ | 0.775 | 0.761 | 0.748 | 0.734 | 4 | 0.409 | 0.411 | 0.386 | 0.365 | 0.359 |
| 50.796 | 0.782 | 0.767 | 0.752 | 0.737 | 5 | 0.432 | 0.427 | 0.406 | 0.378 | 0.359 |
| $S=\Sigma\|\bar{P}(t)-T L\|$ |  |  |  |  |  |  |  |  |  |  |
| 143.96 | 51.32 | 54.40 | 55.23 | 54.52 | 1 | 65.73 | 48.59 | 40.59 | 36.57 | 34.87 |
| 253.23 | 53.94 | 54.34 | 54.37 | 54.14 | 2 | 31.51 | 29.72 | 27.92 | 26.81 | 26.44 |
| 356.79 | 55.75 | 55.29 | 55.04 | 54.87 | 3 | 13.53 | 18.04 | 20.39 | 21.68 | 22.46 |
| 459.77 | 57.88 | 56.81 | 56.17 | 55.76 | 4 | 3.50 | 7.96 | 12.46 | 15.36 | 17.29 |
| $5 \quad 62.58$ | 60.11 | 58.57 | 57.57 | 56.90 | 5 | 13.99 | 5.52 | 6.19 | 9.32 | 11.78 |
| YMEAN $\times 100$ |  |  |  |  |  |  |  |  |  |  |
| 11.49 | 1.27 | 1.19 | 1.20 | 1.26 | 1 | 1.59 | 1.60 | 1.66 | 1.79 | 1.93 |
| 21.36 | 1.33 | 1.32 | 1.33 | 1.33 | 2 | 2.01 | 2.00 | 2.01 | 2.04 | 2.06 |
| $\begin{array}{ll}3 & 1.27\end{array}$ | 1.30 | 1.31 | 1.32 | 1.32 | 3 | 2.12 | 2.11 | 2.10 | 2.04 | 2.06 |
| $4 \quad 1.18$ | 1.24 | 1.27 | 1.29 | 1.30 | 4 | 2.10 | 2.13 | 2.13 | 2.13 | 2.12 |
| 51.08 | 1.17 | 1.22 | 1.25 | 1.27 | 5 | 2.02 | 2.11 | 2.13 | 2.14 | 2.13 |

Table 15
Results for the case where two stocks coexist in one management area. Stock B has suffered high fishing intensity. TL is set at 0.8 . CPUE is proportional to population

| $\mathrm{PT}(\mathrm{N}), \mathrm{CPUE} \propto \mathrm{P}_{\mathrm{t}}, \quad \mathrm{TL}=0.8$ |  |  |  |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Stock A |  |  |  |  |  | Stock B |  |  |  |  |  |
| g/h | 0.02 | 0.04 | 0.06 | 0.08 | 0.10 | g/h | 0.02 | 0.04 | 0.06 | 0.08 | 0.10 |
| PMEAN |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 0.865 | 0.901 | 0.916 | 0.923 | 0.926 | 1 | 0.416 | 0.572 | 0.643 | 0.679 | 0.698 |
| 2 | 0.894 | 0.909 | 0.917 | 0.922 | 0.925 | 2 | 0.574 | 0.636 | 0.672 | 0.692 | 0.704 |
| 3 | 0.909 | 0.917 | 0.922 | 0.925 | 0.927 | 3 | 0.645 | 0.676 | 0.696 | 0.708 | 0.716 |
| 4 | 0.921 | 0.926 | 0.928 | 0.930 | 0.930 | 4 | 0.696 | 0.711 | 0.721 | 0.728 | 0.733 |
| 5 | 0.933 | 0.935 | 0.936 | 0.936 | 0.936 | 5 | 0.738 | 0.744 | 0.748 | 0.750 | 0.752 |
| Minimum population |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 0.711 | 0.727 | 0.739 | 0.742 | 0.743 | 1 | 0.212 | 0.305 | 0.347 | 0.374 | 0.383 |
| 2 | 0.766 | 0.774 | 0.776 | 0.778 | 0.779 | 2 | 0.352 | 0.402 | 0.419 | 0.434 | 0.449 |
| 3 | 0.792 | 0.798 | 0.802 | 0.803 | 0.802 | 3 | 0.420 | 0.450 | 0.463 | 0.476 | 0.487 |
| 4 | 0.806 | 0.810 | 0.811 | 0.810 | 0.809 | 4 | 0.455 | 0.486 | 0.496 | 0.505 | 0.515 |
| 5 | 0.813 | 0.817 | 0.817 | 0.817 | 0.816 | 5 | 0.474 | 0.507 | 0.521 | 0.529 | 0.536 |
| $\boldsymbol{S}=\boldsymbol{\Sigma}\|\bar{P}(t)-T L\|$ |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 16.20 | 21.65 | 24.04 | 25.19 | 25.64 | 1 | 77.44 | 47.79 | 37.29 | 31.40 | 27.79 |
| 2 | 18.76 | 21.80 | 23.47 | 24.41 | 24.92 | 2 | 45.88 | 33.44 | 26.41 | 22.33 | 19.92 |
| 3 | 21.71 | 23.36 | 24.87 | 25.00 | 23.39 | 3 | 31.84 | 25.55 | 21.60 | 19.08 | 17.46 |
| 4 | 24.27 | 25.12 | 25.66 | 26.01 | 26.22 | 4 | 21.61 | 18.51 | 16.45 | 15.09 | 14.20 |
| 5 | 26.64 | 26.97 | 27.16 | 27.26 | 27.30 | 5 | 13.43 | 12.20 | 11.32 | 10.71 | 10.39 |
| YMEAN $\times 100$ |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 1.27 | 1.06 | 0.97 | 0.94 | 0.94 | 1 | 1.73 | 1.72 | 1.70 | 1.71 | 1.74 |
| 2 | 1.18 | 1.08 | 1.02 | 1.00 | 0.98 | 2 | 2.00 | 1.94 | 1.90 | 1.89 | 1.88 |
| 3 | 1.10 | 1.04 | 1.01 | 0.98 | 0.97 | 3 | 2.00 | 1.95 | 1.92 | 1.90 | 1.89 |
| 4 | 1.01 | 0.98 | 0.96 | 0.95 | 0.94 | 4 | 1.93 | 1.89 | 1.88 | 1.87 | 1.86 |
| 5 | 0.92 | 0.91 | 0.91 | 0.90 | 0.90 | 5 | 1.81 | 1.80 | 1.80 | 1.80 | 1.81 |

The following notation is used:
$\mathrm{P}_{\mathrm{A}, \mathrm{t}}=$ relative size of stock A at the beginning of year t
$\mathrm{P}_{\mathrm{B}, \mathrm{t}}=$ relative size of B at the beginning of year t
$\mathrm{Y}_{\mathrm{A}, \mathrm{t}}=$ relative catch from stock A in year t
$Y_{B, t}=$ relative catch from stock $B$ in year $t$
$Y_{t}=$ total catch in the area in year $t$
$X_{t}=$ CPUE in the area in year $t$
$Y_{A, t}, Y_{B, t}$ are calculated from the equations

$$
\begin{align*}
& Y_{A, t}=\alpha_{t} Y_{t}  \tag{22}\\
& Y_{B, t}=\left(1-\alpha_{t}\right) Y_{t} \tag{23}
\end{align*}
$$

Where $\alpha_{t}$ is decided from the equation

$$
\begin{equation*}
\alpha_{\mathrm{t}}=\alpha_{0} \cdot\left(\sqrt{\mathrm{P}_{\mathrm{A}, \mathrm{t}}} / \sqrt{\mathrm{P}_{\mathrm{B}, \mathrm{t}}}\right) \cdot\left(1+\gamma_{\mathrm{t}}\right) \tag{24}
\end{equation*}
$$

$\alpha_{0}\left(0 \leq \alpha_{0} \leq 1\right)$ is set at 0.3 assuming a high fishing intensity for stock B. $\alpha_{t}$ is considered to change according to the ratio of stock size A to B . It is assumed that changes are proportional to the ratio of the square root of each population (Table 13). $\gamma_{t}$ is added as a random fluctuation term of fishing intensity, which uniformly fluctuates from -0.5 to 0.5 .
Reproduction takes place independently in each stock:

$$
\begin{align*}
& \mathbf{P}_{\mathrm{A}, \mathrm{t}+1}=\left(\mathrm{P}_{\mathrm{A}, \mathrm{t}}-\mathrm{Y}_{\mathrm{A}, \mathrm{t}}\right) \mathrm{e}^{-\mathrm{M}}+\mathrm{M}[1+\mathrm{A}\{1- \\
& \left.\left.\left(\mathrm{P}_{\mathrm{A}, \mathrm{t}}-1\right)^{\mathrm{n}}\right\}\right] \mathrm{P}_{\mathrm{A}, \mathrm{t}}-10  \tag{25}\\
& \mathbf{P}_{\mathrm{B}, \mathrm{t}+1}=\left(\mathbf{P}_{\mathrm{B}, \mathrm{t}}-\mathrm{Y}_{\mathrm{B}, \mathrm{t}}\right) \mathrm{e}^{-\mathrm{M}}+\mathrm{M}[1+\mathrm{A}\{1- \\
& \left.\left.\left.\quad\left(\mathrm{P}_{\mathrm{B}, \mathrm{t}-l}\right)\right)^{\mathrm{n}}\right\}\right] \mathrm{P}_{\mathrm{B}, \mathrm{t}}-l \tag{26}
\end{align*}
$$

The observed CPUE in the total area is given from either of the following equations, with weight $\alpha_{t}$ :

$$
\begin{align*}
& X_{t}=k_{1}\left[\alpha_{t} P_{A, t}+\left(1-\alpha_{t}\right) P_{B, t}\right]\left(1+\varepsilon_{t}^{\prime}\right)  \tag{27}\\
& X_{t}=k_{2} \sqrt{\alpha_{t} P_{A, t}+\left(1-\alpha_{t}\right) P_{B, t}}\left(1+\varepsilon_{t}^{\prime}\right) \tag{28}
\end{align*}
$$

where, $\varepsilon_{\mathrm{t}}$, is the observation error randomly fluctuating uniformly from -0.35 to 0.35 . For simplicity, we set $k_{1}=k_{2}=1$. The decision process for the catch quota follows equations (4), (5) and (6).
The results of the simulation runs are shown in Tables 14-17 and some typical examples illustrated in Figs 6 and 7.


Fig. 6. Population trajectories of stock $A+B$, stock $A$ and stock $B$ respectively, in the case where CPUE is proportional to population. $\mathrm{g}=3, \mathrm{~h}=0.02$.


Fig. 7. Series of catch quotas and observed CPUE values for the case where two stocks coexist in one area. CPUE is proportional to population. $\mathrm{g}=3, \mathrm{~h}=0.02$.

Tables 14 and 15, and Figs 6 and 7 show that when the CPUE is proportional to the population level, extinction does not occur in either stock. Stock A (low fishing intensity) stabilises at a higher level than TL while Stock B (high fishing intensity) stabilises at a lower level than TL.

For the case where CPUE is proportional to the square root of population size (Tables 16 and 17) extinction occurs in both stocks when TL is set at 0.6 and g is set at 1.0 . To avoid this $g$ must be set higher than 1.0 or the TL should be set higher than the MSYL.

## One stock distributed over two management areas

This section considers the case where one biological stock is found in two management areas (Areas I and II). The stability of the system is examined for the case where the management procedure is applied treating the two Areas as independent stocks.

The following notation is used:
$P_{t}=$ total stock size at the beginning of year $t$
$\mathrm{P}_{\mathrm{I}, \mathrm{t}}=$ number of whales migrating into Area I at the beginning of year $t$
$\mathrm{P}_{\mathrm{II}, \mathrm{t}}=$ number of whales migrating into Area II at the beginning of year $t$
$\mathrm{X}_{\mathrm{I}, \mathrm{t}}=$ CPUE in Area I in year t
$\mathrm{X}_{\mathrm{II}, \mathrm{t}}=$ CPUE in Area II in year t
$P_{I, t}$ and $P_{I I, t}$ are calculated from the equations:

$$
\begin{gather*}
\mathrm{P}_{\mathrm{I}, \mathrm{t}}=\beta_{0}\left(1+\gamma_{\mathrm{t}}\right) \mathrm{P}_{\mathrm{t}}  \tag{29}\\
\mathrm{P}_{\mathrm{II}, \mathrm{t}}=\mathrm{P}_{\mathrm{t}}-\mathrm{P}_{\mathrm{I}, \mathrm{t}} \tag{30}
\end{gather*}
$$

where $\beta_{0}$ is the mean ratio of the population migrating into the Area I to the whole population and $\gamma_{t}$ is a random variable which uniformly fluctuates from -0.5 to 0.5 . It is assumed that $\beta_{0}$ is independent of the level of $P_{t}$.
CPUE is independently observed in each Area. Two cases are assumed, that is:

$$
\begin{equation*}
\mathrm{X}_{\mathrm{I}, \mathrm{t}}=\mathrm{k}_{1} \mathrm{P}_{\mathrm{I}, \mathrm{t}}, \quad \mathrm{X}_{\mathrm{II}, \mathrm{t}}=\mathrm{k}_{2} \mathrm{P}_{\mathrm{II}, \mathrm{t}} \tag{31}
\end{equation*}
$$

or

$$
\begin{equation*}
\mathrm{X}_{\mathrm{I}, \mathrm{t}}=\mathrm{k}_{1} \sqrt{\mathrm{P}_{\mathrm{I}, \mathrm{t}}}, \quad \mathrm{X}_{\mathrm{II}, \mathrm{t}}=\mathrm{k}_{2} \sqrt{\mathrm{P}_{\mathrm{II}, \mathrm{t}}} \tag{32}
\end{equation*}
$$

where $k_{1}$ and $k_{2}$ are proportional constants. For simplicity we set $\mathrm{k}_{1}=\mathrm{k}_{2}=1$.
The catch quota is calculated as follows:

$$
\begin{array}{ll}
\mathrm{Y}_{\mathrm{i}, \mathrm{t}+1}=\left(1+\mathrm{hL}_{\mathrm{i}, \mathrm{t}-l}+\mathrm{gK}_{\mathrm{i}, \mathrm{t}-l}\right) \mathrm{Y}_{\mathrm{i}, \mathrm{t}}, & \mathrm{i}=\mathrm{I}, \mathrm{II} \\
\mathrm{~L}_{\mathrm{i}, \mathrm{t}-l}=\mathrm{E}\left(\mathrm{X}_{\mathrm{i}, \mathrm{t}-l}\right) / \mathrm{TL}_{\mathrm{i}-1} & \mathrm{i}=\mathrm{I}, \mathrm{II} \\
\mathrm{~K}_{\mathrm{i}, \mathrm{t}-l}=\mathrm{b}_{\mathrm{i}} / \mathrm{E}\left(\mathrm{X}_{\mathrm{i}, \mathrm{t}-l}\right), & \mathrm{i}=\mathrm{I}, \mathrm{II}
\end{array}
$$

## Table 16

Results for the case where two stocks coexist in one management area. Stock B has suffered high fishing intensity. TL is set at 0.6. CPUE is proportional to the square root of population

| PT(N), CPUE $\propto \sqrt{ } \mathrm{P}_{\mathrm{t}}, \quad \mathrm{TL}=0.6$ |  |  |  |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Stock A |  |  |  |  |  | Stock B |  |  |  |  |  |
|  | 0.02 | 0.04 | 0.06 | 0.08 | 0.10 | g/h | 0.02 | 0.04 | 0.06 | 0.08 | 0.10 |
| PMEAN |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 0.264 | 0.278 | 0.316 | 0.328 | 0.334 | 1 | 0.155 | 0.107 | 0.104 | 0.102 | 0.097 |
| 2 | 0.687 | 0.687 | 0.702 | 0.718 | 0.729 | 2 | 0.192 | 0.178 | 0.176 | 0.182 | 0.192 |
| 3 | 0.793 | 0.769 | 0.763 | 0.764 | 0.748 | 3 | 0.284 | 0.242 | 0.225 | 0.219 | 0.219 |
| 4 | 0.839 | 0.812 | 0.797 | 0.789 | 0.786 | 4 | 0.384 | 0.312 | 0.276 | 0.257 | 0.247 |
| 5 | 0.865 | 0.838 | 0.821 | 0.810 | 0.802 | 5 | 0.476 | 0.385 | 0.332 | 0.302 | 0.283 |
| Minimum population |  |  |  |  |  |  |  |  |  |  |  |
| 1 | - | - | - |  |  | 1 | - | - |  | - | - |
| 2 | 0.448 | 0.453 | 0.385 | 0.312 | 0.246 | 2 | 0.038 | 0.039 | 0.033 | 0.026 | 0.019 |
| 3 | 0.600 | 0.567 | 0.506 | 0.442 | 0.377 | 3 | 0.083 | 0.079 | 0.072 | 0.076 | 0.063 |
| 4 | 0.644 | 0.594 | 0.542 | 0.487 | 0.433 | 4 | 0.142 | 0.108 | 0.092 | 0.083 | 0.083 |
| 5 | 0.650 | 0.605 | 0.556 | 0.503 | 0.448 | 5 | 0.198 | 0.147 | 0.114 | 0.094 | 0.085 |
| $S=\Sigma\|\bar{P}(t)-T L\|$ |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 81.45 | 77.37 | 69.03 | 65.90 | 64.66 | 1 | 100.85 | 102.51 | 103.11 | 103.65 | 104.56 |
| 2 | 17.44 | 21.95 | 30.04 | 37.09 | 42.72 | 2 | 85.48 | 88.40 | 88.70 | 87.50 | 85.48 |
| 3 | 38.59 | 33.90 | 32.56 | 62.79 | 34.46 | 3 | 67.18 | 75.55 | 79.05 | 80.27 | 80.23 |
| 4 | 47.72 | 12.39 | 39.36 | 37.85 | 37.16 | 4 | 47.22 | 61.61 | 68.75 | 72.51 | 74.53 |
| 5 | 53.02 | 47.68 | 42.20 | 41.96 | 40.46 | 5 | 28.78 | 47.03 | 57.50 | 63.66 | 67.46 |
| YMEAN |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 1.54 | 1.42 | 1.29 | 1.20 | 1.14 | 1 | 1.17 | 1.12 | 1.09 | 1.06 | 1.04 |
| 2 | 1.94 | 1.81 | 1.65 | 1.50 | 1.37 | 2 | 1.41 | 1.32 | 1.27 | 1.24 | 1.22 |
| 3 | 1.70 | 1.75 | 1.72 | 1.67 | 1.61 | 3 | 1.69 | 1.55 | 1.47 | 1.43 | 1.41 |
| 4 | 1.50 | 1.62 | 1.66 | 1.66 | 1.65 | 4 | 1.91 | 1.76 | 1.65 | 1.58 | 1.54 |
| 5 | 1.37 | 1.51 | 1.58 | 1.61 | 1.63 | 5 | 2.03 | 1.93 | 1.81 | 1.72 | 1.66 |

Table 17
Results for the case where two stocks coexist in one management area. Stock B has suffered high fishing intensity. TL is set at 0.8 . CPUE is proportional to the square root of population

| PT(N), CPUE $/ \mathrm{P}_{\mathrm{t}}, \quad \mathrm{TL}=0.8$ |  |  |  |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Stock A |  |  |  |  |  | Stock B |  |  |  |  |  |
| g/h | 0.02 | 0.04 | 0.06 | 0.08 | 0.10 | $\mathrm{g} / \mathrm{h}$ | 0.02 | 0.04 | 0.06 | 0.08 | 0.10 |
| PMEAN |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 0.578 | 0.744 | 0.807 | 0.840 | 0.860 | 1 | 0.156 | 0.205 | 0.270 | 0.340 | 0.400 |
| 2 | 0.806 | 0.842 | 0.861 | 0.873 | 0.881 | 2 | 0.285 | 0.343 | 0.400 | 0.446 | 0.481 |
| 3 | 0.858 | 0.869 | 0.876 | 0.880 | 0.883 | 3 | 0.425 | 0.455 | 0.481 | 0.501 | 0.516 |
| 4 | 0.883 | 0.885 | 0.886 | 0.887 | 0.887 | 4 | 0.529 | 0.536 | 0.541 | 0.546 | 0.549 |
| 5 | 0.901 | 0.889 | 0.897 | 0.895 | 0.894 | 5 | 0.607 | 0.600 | 0.594 | 0.589 | 0.585 |
| Minimum population |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 0.249 | 0.467 | 0.505 | 0.528 | 0.531 | 1 | 0.017 | 0.038 | 0.056 | 0.071 | 0.082 |
| 2 | 0.586 | 0.635 | 0.632 | 0.630 | 0.626 | 2 | 0.100 | 0.137 | 0.152 | 0.161 | 0.171 |
| 3 | 0.670 | 0.664 | 0.659 | 0.654 | 0.648 | 3 | 0.161 | 0.195 | 0.228 | 0.246 | 0.246 |
| 4 | 0.684 | 0.679 | 0.673 | 0.667 | 0.661 | 4 | 0.208 | 0.276 | 0.303 | 0.313 | 0.309 |
| 5 | 0.686 | 0.681 | 0.675 | 0.669 | 0.664 | 5 | 0.248 | 0.276 | 0.303 | 0.313 | 0.309 |
| $S=\Sigma\|\vec{P}(t)-T L\|$ |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 47.27 | 27.40 | 26.69 | 27.54 | 28.40 | 1 | 129.57 | 119.81 | 106.69 | 92.87 | 81.21 |
| 2 | 10.86 | 15.26 | 18.21 | 20.10 | 21.29 | 2 | 103.73 | 92.14 | 80.77 | 71.44 | 64.48 |
| 3 | 12.26 | 14.54 | 16.05 | 17.07 | 17.75 | 3 | 75.81 | 69.67 | 64.50 | 60.47 | 57.48 |
| 4 | 16.57 | 16.97 | 17.22 | 17.37 | 17.44 | 4 | 54.90 | 53.59 | 52.48 | 51.60 | 50.95 |
| 5 | 20.25 | 19.81 | 19.42 | 19.08 | 18.9 | 5 | 39.29 | 40.67 | 41.87 | 42.89 | 43.75 |
| YMEAN $\times 100$ |  |  |  |  |  |  |  |  |  |  |  |
|  | 1.87 | 1.54 | 1.31 | 1.16 | 1.07 | 1 | 1.28 | 1.31 | 1.32 | 1.32 | 1.32 |
| 2 | 1.58 | 1.38 | 1.26 | 1.19 | 1.16 | 2 | 1.62 | 1.64 | 1.66 | 1.67 | 1.70 |
| 3 | 1.37 | 1.30 | 1.26 | 1.23 | 1.22 | 3 | 1.89 | 1.89 | 1.89 | 1.91 | 1.93 |
| 4 | 1.24 | 1.23 | 1.22 | 1.22 | 1.22 | 4 | 1.98 | 1.99 | 2.00 | 2.01 | 2.02 |
| 5 | 1.13 | 1.15 | 1.16 | 1.18 | 1.19 | 5 | 1.95 | 1.99 | 2.01 | 2.03 | 2.04 |

where $b_{i}$ denotes the slope of the regression line fitted to the series of CPUE from $t-l-5$ to $t-l+5$ in Areai. TL and $\mathrm{TL}_{\text {II }}$ denote the target levels for Areas I and II respectively. They are assumed to be set at 0.6 or 0.8 of the initial level of CPUE in each Area, i.e., $\beta_{0} \mathrm{P}_{0},\left(1-\beta_{0}\right) \mathrm{P}_{0}$.

The dynamics of the stock are represented by

$$
\begin{align*}
& \mathrm{P}_{\mathrm{t}+1}=\left(\mathrm{P}_{\mathrm{t}}-\mathrm{Y}_{\mathrm{I}, \mathrm{t}}-\mathrm{Y}_{\mathrm{II}, \mathrm{t}}\right) \mathrm{e}^{-\mathrm{M}} \\
& \quad+\mathrm{M}\left[1+\mathrm{A}\left\{1-\left(\mathrm{P}_{\mathrm{t}-1}\right)^{\mathrm{n}}\right\}\right] \mathrm{P}_{\mathrm{t}-1} \tag{36}
\end{align*}
$$

Table 18
Results for the case where one stock is managed in two independent management areas. CPUE is proportional to population

| $\operatorname{PT}(\mathrm{N}), \mathrm{CPUE} \alpha \mathrm{P}_{t}, \quad \mathrm{TL}=0.6$ |  |  |  |  |  | $\operatorname{PT}(\mathrm{N}), \mathrm{CPUE} \mathrm{\propto P} \mathrm{P}_{\mathrm{t}}, \mathrm{TL}=0.8$ |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathrm{g} / \mathrm{h}$ | 0.02 | 0.04 | 0.06 | 0.08 | 0.10 | g/h | 0.02 | 0.04 | 0.06 | 0.08 | 0.10 |
| PMEAN |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 0.581 | 0.598 | 0.613 | 0.616 | 0.609 | 1 | 0.703 | 0.746 | 0.767 | 0.779 | 0.785 |
| 2 | 0.669 | 0.643 | 0.632 | 0.627 | 0.622 | 2 | 0.759 | 0.775 | 0.785 | 0.791 | 0.794 |
| 3 | 0.726 | 0.686 | 0.664 | 0.651 | 0.643 | 3 | 0.798 | 0.803 | 0.806 | 0.807 | 0.808 |
| 4 | 0.771 | 0.726 | 0.698 | 0.979 | 0.666 | 4 | 0.831 | 0.830 | 0.829 | 0.827 | 0.825 |
| 5 | 0.810 | 0.764 | 0.731 | 0.708 | 0.692 | 5 | 0.860 | 0.856 | 0.851 | 0.847 | 0.843 |
| Minimum population |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 0.443 | 0.419 | 0.402 | 0.369 | 0.341 | 1 | 0.574 | 0.602 | 0.608 | 0.611 | 0.612 |
| 2 | 0.569 | 0.532 | 0.512 | 0.502 | 0.483 | 2 | 0.646 | 0.657 | 0.663 | 0.665 | 0.662 |
| 3 | 0.613 | 0.568 | 0.530 | 0.498 | 0.476 | 3 | 0.672 | 0.679 | 0.683 | 0.686 | 0.688 |
| 4 | 0.623 | 0.566 | 0.529 | 0.508 | 0.495 | 4 | 0.676 | 0.682 | 0.687 | 0.686 | 0.688 |
| 5 | 0.618 | 0.558 | 0.504 | 0.472 | 0.452 | 5 | 0.668 | 0.667 | 0.666 | 0.665 | 0.665 |
| $S=\Sigma\|\bar{P}(t)-T L\|$ |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 14.39 | 17.47 | 20.44 | 21.69 | 24.14 | 1 | 20.81 | 14.63 | 13.22 | 12.25 | 11.35 |
| 2 | 13.88 | 8.58 | 8.62 | 9.28 | 9.45 | 2 | 9.69 | 8.36 | 8.09 | 7.73 | 7.36 |
| 3 | 25.64 | 17.24 | 12.79 | 10.18 | 8.50 | 3 | 6.87 | 6.87 | 6.76 | 6.59 | 6.36 |
| 4 | 34.16 | 25.29 | 19.52 | 15.75 | 13.19 | 4 | 8.87 | 8.55 | 8.18 | 7.79 | 7.40 |
| 5 | 41.83 | 32.84 | 26.30 | 21.69 | 18.38 | 5 | 13.32 | 12.45 | 11.60 | 10.79 | 10.05 |
| YMEAN $\times 100$ |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 2.06 | 1.99 | 2.00 | 2.06 | 2.10 | 1 | 1.91 | 1.80 | 1.74 | 1.72 | 1.71 |
| 2 | 2.07 | 2.09 | 2.10 | 2.11 | 2.11 | 2 | 1.81 | 1.75 | 1.72 | 1.71 | 1.70 |
| 3 | 1.96 | 2.05 | 2.08 | 2.09 | 2.10 | 3 | 1.67 | 1.65 | 1.65 | 1.65 | 1.65 |
| 4 | 1.81 | 1.96 | 2.03 | 2.06 | 2.08 | 4 | 1.50 | 1.52 | 1.54 | 1.55 | 1.57 |
| 5 | 1.63 | 1.84 | 1.94 | 2.00 | 2.03 | 5 | 1.33 | 1.37 | 1.41 | 1.44 | 1.47 |

Table 19
Results for the case where one stock is managed in two independent management areas. CPUE is proportional to the square root of population

| $\operatorname{PT}(\mathrm{N}), \mathrm{CPUE} \propto \sqrt{ } \mathrm{P}_{\mathrm{t}}, \quad$ TL=0.6 |  |  |  |  | PT(N), CPUE $\propto / \mathrm{P}_{\mathrm{t}}, \quad \mathrm{TL}=0.8$ |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathrm{g} / \mathrm{h} 0.02$ | 0.04 | 0.06 | 0.08 | 0.10 | g/h | 0.02 | 0.04 | 0.06 | 0.08 | 0.10 |
| PMEAN |  |  |  |  |  |  |  |  |  |  |
| 10.158 | 0.122 | 0.106 | 0.085 | 0.809 | 1 | 0.203 | 0.158 | 0.135 | 0.120 | 0.110 |
| 20.219 | 0.154 | 0.118 | 0.099 | 0.091 | 2 | 0.354 | 0.365 | 0.341 | 0.224 | 0.189 |
| 30.323 | 0.235 | 0.214 | 0.209 | 0.183 | 3 | 0.464 | 0.354 | 0.313 | 0.300 | 0.299 |
| 40.405 | 0.285 | 0.239 | 0.223 | 0.212 | 4 | 0.545 | 0.416 | 0.354 | 0.323 | 0.306 |
| 50.476 | 0.101 | 0.273 | 0.244 | 0.237 | 5 | 0.612 | 0.474 | 0.399 | 0.357 | 0.331 |
| Minimum population |  |  |  |  |  |  |  |  |  |  |
| 1 - |  |  |  | - | 1 | - | - | - |  |  |
| 20.012 | - | - |  | - | 2 | 0.102 | 0.034 | 0.023 | 0.008 |  |
| 30.101 | 0.040 | 0.036 |  | - | 3 | 0.226 | 0.110 | 0.073 | 0.071 | 0.159 |
| 40.156 | 0.079 | 0.056 |  | - | 4 | 0.309 | 0.173 | 0.116 | 0.094 | 0.088 |
| 50.209 | 0.108 | 0.075 | 0.03 | - | 5 | 0.378 | 0.205 | 0.158 | 0.116 | 0.091 |
| $S=\Sigma\|\bar{P}(t)-T L\|$ |  |  |  |  |  |  |  |  |  |  |
| 196.91 | 103.48 | 106.52 | 108.34 | 109.53 | 1 | 120.82 | 129.73 | 134.38 | 137.31 | 139.32 |
| 284.90 | 97.20 | 104.11 | 107.52 | 109.10 | 2 | 92.40 | 108.28 | 113.13 | 116.63 | 123.57 |
| 364.09 | 81.10 | 84.79 | 85.60 | 90.65 | 3 | 68.67 | 90.63 | 98.84 | 101.35 | 101.67 |
| 447.97 | 71.10 | 79.81 | 82.72 | 84.76 | 4 | 52.37 | 78.11 | 90.51 | 96.75 | 100.25 |
| 534.18 | 61.40 | 73.07 | 78.63 | 79.81 | 5 | 39.02 | 66.59 | 81.67 | 90.08 | 95.15 |
| YMEAN $\times 100$ |  |  |  |  |  |  |  |  |  |  |
| 11.27 | 1.19 | 1.20 | 1.24 | 1.30 | 1 | 1.41 | 1.26 | 1.20 | 1.17 | 1.17 |
| 21.44 | 1.19 | 1.10 | 1.09 | 1.11 | 2 | 1.81 | 1.49 | 1.30 | 1.19 | 1.12 |
| 31.76 | 1.44 | 1.34 | 1.25 | 1.18 | 3 | 2.07 | 1.80 | 1.64 | 1.58 | 1.60 |
| $4 \quad 1.97$ | 1.63 | 1.48 | 1.42 | 1.36 | 4 | 2.15 | 1.97 | 1.80 | 1.71 | 1.66 |
| $5 \quad 2.09$ | 1.78 | 1.60 | 1.51 | 1.41 | 5 | 2.14 | 2.08 | 1.92 | 1.81 | 1.74 |

The results are shown in Tables 18 and 19 and Figs 8 and 9. Table 18 and Figs 8 and 9 are for the case where CPUE is proportional to population size. For the case assumed here, no serious problems occurred. However, where CPUE is proportional to the square root of the population size, many cases of extinction appeared. To avoid this, TL must be set above MSYL and a large $g$ value should be taken.


Fig. 8. Population trajectories for the total stock, and those migrating into Area I and II. CPUE is proportional to population. $\mathrm{g}=3, \mathrm{~h}$ $=0.02$.


Fig. 9. Series of catch quotas, and observed CPUE values for Area I and II. CPUE is proportional to population. $\mathrm{g}=3, \mathrm{~h}=0.02$.

## DISCUSSION

For the purposes of these simulations, during the initial ten years a constant catch of $4 \%$ of the initial stock size was assumed. However, the initial stock size is not usually known exactly. To examine this we ran the simulations for constant catches of $2,4,6$ or $8 \%$ of the initial stock size. The results are essentially the same as for $4 \%$, showing that in practice this assumption does not create a serious problem for the management procedure.
The aim of these simulations is to verify that successful management can be carried out even if little information on the stock is available. Therefore, no attempt was made to estimate the reproduction curve, the MSY or MSYL.

However, if needed, a procedure could be added which would estimate MSY or MSYL using cumulative pairs of CPUE and $Y_{t}$ data.

In these simulations, the mean and trend of CPUE over 11 years were used to regulate the catch quota. This seems to be an important way to reduce the effect of noise in the CPUE. The procedure also incorporates a feedback system for adjusting $g$ depending upon the predicted CPUE value, $\hat{\mathbf{X}}_{\mathrm{t}}$. This should reduce the instability of the system due to time lags as discussed by Tanaka (1980).

In this simulation, the calculation of the discriminant function or auto-regression model was done for every 10 years to save time. By shortening this period, a more sensitive feedback control could be achieved.

It should be noted that the ways of modifiying the set TL or $g$ value shown here are only examples and more suitable ways could be designed.

Although the detailed results are not presented here, further simulations were conducted with changing mean values for the Poisson distribution for the frequency distribution of the 'perturbation'. The results suggest that as would be expected the effects of the disturbance depend on how large and/or how often the disturbance occurs.

The simulations reported here used the Southern Hemisphere minke whale populations as a model. However, the management procedure could be applied to other whale or fish stocks with only small modifications.

## SUMMARY

Some combinations of $g$ and $h$ values give high stability to the system and the stock is maintained around the target level. The stability of the system greatly depends on whether or not supercompensation is assumed.

Whether or not the index of abundance is directly proportional to population size is an important factor.

Where CPUE is proportional to the square root of the population size, large values of $g$, smaller values of $h$ and a high target level should be chosen. Although the possibility of stock extinction was negligible if CPUE is proportional to population size, for the case where CPUE is proportional to the square root of population size, caution is needed. In the latter case, a $g$ value greater than 2 and a TL of 0.8 would provide satisfactory results.

Noise in the CPUE of the levels assumed here did not produce serious problems for the procedure. Similarly, the sudden decrease in population assumed here produced no serious problems.

It is very important which values of TL and $g$ are chosen; in some cases, modification of the set TL can be done successfully. Stability was increased by incorporation of the $g$ value into a feedback system.

Robust management can be achieved even in the case where biological stock boundary does not coincide with management area. However, caution is needed for the case where CPUE is proportional to the square root of the population size.

## ACKNOWLEDGEMENTS

The authors are indebted to Dr. Y. Matsumiya, Ocean Research Institute, University of Tokyo, 1-15-1 Minamidai, Nakano-Ku, Tokyo 164, Japan, and Dr. H. Kishino, Institute of Statistical Mathematics, 4-6-7 Minamiazabu, Minato-Ku, Tokyo 106, Japan, and two anonymous reviewers for their useful comments on this manuscript.

## REFERENCES

Akaike, H. 1971. Autoregressive model fitting for control. Ann. Inst. Statist. Math. 23: 168-80.
de la Mare, W.K. 1985a. Fitting population models to time series of abundance data. Rep. int. Whal. Commn 36: 399-418.
de la Mare, W.K. 1985b. Simulation studies of management procedures. Rep. int. Whal. Commn 36: 421-50.
Holt, S.J. 1985. The classification of whale stocks and the determination of catch limits, under the current IWC management procedure, with limited information. Rep. int. Whal. Commn 35: 487-94.
Pella, J.J. and Tomlinson, P.K. 1969. A generalised stock-production model. Bull. I-ATTC 13: 421-96.
Sakuramoto, K. and Tanaka, S. 1986. A simulation study on management whale stocks considering feedback systems. Paper SC/38/O 10 presented to the IWC Scientific Committee, June 1986. (Unpublished) 23pp.
Sakuramoto, K. and Tanaka, S. 1987. Further simulation study on management whale stocks considering feedback systems. Paper SC/M87/1 presented to the Special Meeting of the IWC Scientific Committee, March 1987. (Unpublished) 18pp.
Shepherd, J.G. 1982. A family of general production curves for exploited populations. Math. Biosci. 59: 77-93.
Tanaka, S. 1980. A theoretical consideration on the management of a stock-fishery system by catch quota and on its dynamical properties. Bull. Japan. Soc. Fish. 46: 1477-82.
Tanaka, S. 1984. On the method for calculating catch quotas. Bull. INPFC 42: 98-103.
Tanaka, S. 1986. On the practical method for stock management. Paper SC/A86/CA6 presented to the Special Meeting of the IWC Scientific Committee, April 1986 (Unpublished) 4pp.


[^0]:    ${ }^{1}$ The Schedule contains the IWC Regulations concerning whaling operations (definitions, catch limits, seasons etc). To amend a provision of the Schedule a three-quarters majority of the votes cast is required.
    ${ }_{2}$ Detailed accounts of the discussions held by the Commission and the Scientific Committee can be found in the Reports of the International Whaling Commission, published annually by the Commission.

[^1]:    ${ }^{1}$ Pollock, K. H. 1987. Mark and recapture techniques for estimating cetacean abundance. Paper SC/39/O 5 presented to the IWC Scientific Committee, June 1987. Unpublished but available from the office of the Commission.

[^2]:    ${ }^{1}$ e.g. IWC, 1987. International Convention for the Regulation of Whaling, 1946. Schedule. International Whaling Commission, Cambridge. 27pp.

[^3]:    * This was provided and is published as Sigurjónsson, J. 1988. Operational factors of the Icelandic large whale fishery. Rep. int. Whal. Commn 38.

[^4]:    * This has been modified in accordance with the Scientific Committee's recommendations at the 1987 Annual Meeting.

[^5]:    * Originally presented as paper SC/39/O 8 to the IWC Scientific Committee.

[^6]:    * These intervals are combined in Fig. 3.
    ** These intervals are combined in Fig. 6.

[^7]:    * Further information can be obtained from R.B. Selander at the Department of Biology, University of Rochester, NY 14627, USA.

[^8]:    * The data are taken from Jeffreys et al. (1986) and further citation or use of these data should refer directly to that publication.

[^9]:    * Originally presented as paper SC/40/Mi6 to the IWC Scientific Committee.

[^10]:    ${ }^{1}$ Two marks recovered from a single carcase. Recorded as separate whales when marked. Brown (IWC, 1984: 95) noted that they were recorded as the same whale when marked. ${ }^{2}$ Same whale. ${ }^{3}$ Discrepancy in records as to whether these marks were fired at one whale or three. ${ }^{4}$ Recorded as different whales when marked; 31466 was found during meat processing and could not be assigned to carcase. ${ }^{5}$ Two marks recovered from a single carcase. Recorded as separate whales when marked. 637007 and 37010 were not assigned to carcase. Recorded as separate whales when marked ${ }^{7}$ Found in meat 7 days apart. ${ }^{8} 38551$ found in meat. Recorded as separate whales when marked. ${ }^{9} 35425$ found in meat. Recorded as the same whale when marked. ${ }^{10}$ Presumably an error, possibly in recording the mark number. This recovery is excluded from analyses.

[^11]:    * Orignally presented as Paper $\mathrm{SC} / 38 / \mathrm{O} 3$ to the IWC Scientific Committee.

