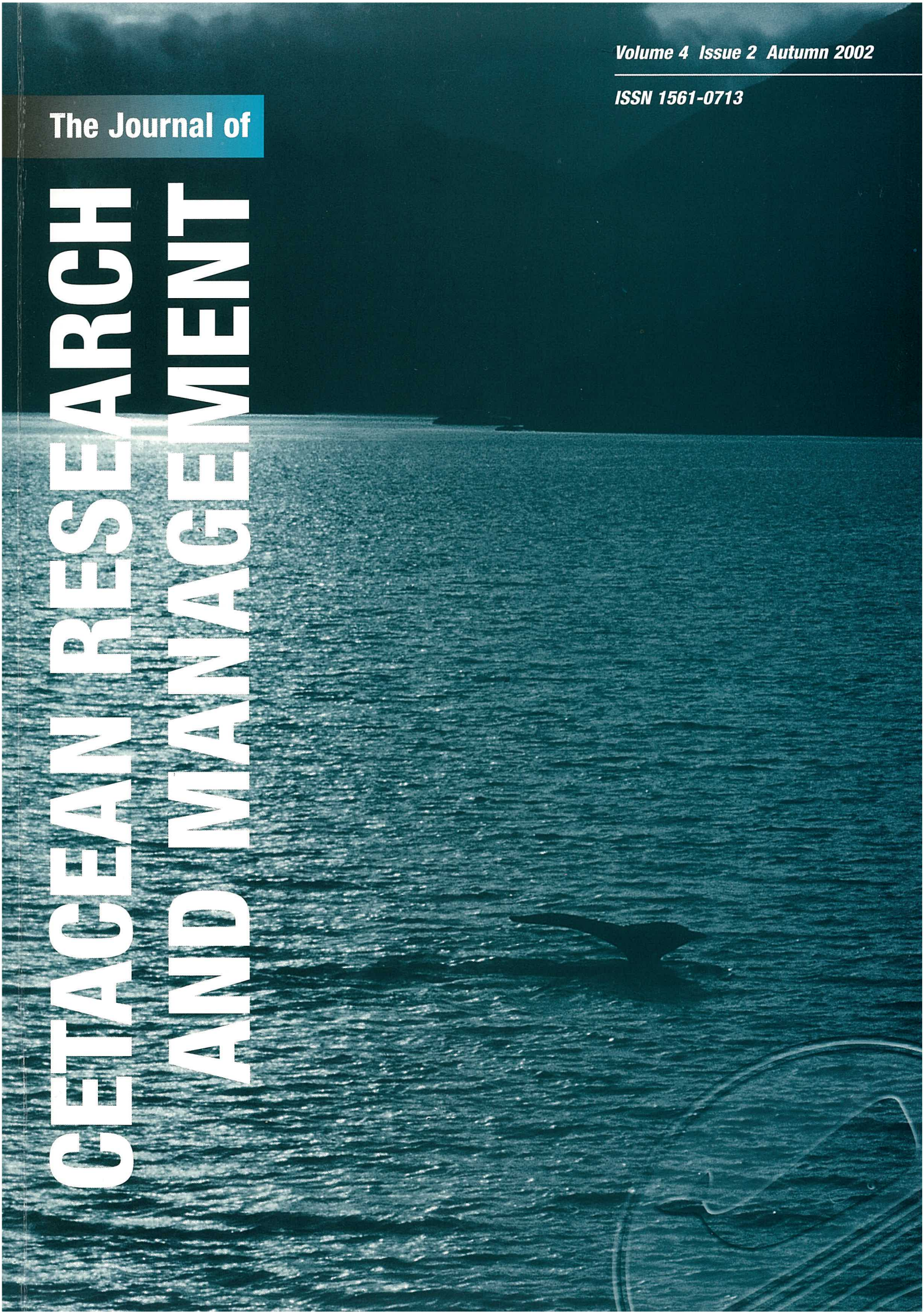
Volume 4 Issue 2 Autumn 2002

ISSN 1561-0713

The Journal of CETACEAN RESEARCH AND MANAGEMENT



Contents

Editorial

This issue of the Journal follows the 2002 meeting of the International Whaling Commission held in Shimonoseki, Japan. Details of the Commission meeting will be published in the next *Annual Report of the International Whaling Commission*. The full report of the Scientific Committee will be published in spring 2003 as *J. Cetacean Res. Manage*. 5 (Suppl.). However, it seems timely to provide a short summary of the work of the Scientific Committee that updates the summary provided in Donovan (2002).

REVISED MANAGEMENT PROCEDURE

After the adoption of the moratorium on commercial whaling in 1982, the Committee spent over eight years developing the Revised Management Procedure (RMP) for baleen whales (IWC, 1999b). In brief, the RMP is a generic management procedure designed to estimate safe catch limits for commercial whaling of baleen whales. This was adopted some time ago by the Commission (IWC, 1993). However, the Commission has stated that it will not set catch limits for commercial whaling for any stocks until it has agreed and adopted a complete Revised Management Scheme (RMS). The RMS will also include a number of non-scientific matters, including inspection and enforcement. This is the subject of a considerable amount of discussion within the Commission, which is holding a special meeting of Commissioners to address this issue in October 2002.

Implementation Simulation Trials

Implementation Simulation Trials are trials that are carried out before using the RMP to calculate a catch limit and involve investigating the full range of plausible hypotheses related to a specific species and geographic area.

The process of developing *Implementation Simulation Trials* is not the same as identifying the 'best' assessment for the species/region, but involves considering a set of alternative models to examine a broad range of uncertainties with a view to excluding variants of the RMP that show performance that is not sufficiently robust across the trials. Account needs to be taken of the plausibility of the various trial scenarios when evaluating RMP variants.

The Committee discussed the general question of how best to ensure that the process of carrying out *Implementations* (or *Implementation Reviews*) is efficient and prompt, whilst taking into account the available information. To achieve this it agreed that they should be conducted at discrete intervals, using the data available at one point in time. The process from *'pre-Implementation Assessment'* to initial *Implementation* and *Implementation Reviews* has been formalised and clarified.

North Pacific common minke whales

The Committee has been working on *Implementation Simulation Trials* for this area since 1994; a special workshop was held prior to the Shimonoseki meeting. It is proving to be difficult for a number of reasons, including: (1) harvesting is projected to take place on migration as well as on feeding grounds; (2) there is a seasonally-dependent overlap of management stocks; (3) continual updating of information on relatively complex population structure; (4) issues related to the plausibility of trials, particularly with respect to population structure; (5) complexity and time required to code and run trials; (6) lack of agreement on when to stop 'improving'. An ambitious workplan has been established with the aim of reaching agreement on the appropriate variant of the RMP to apply to common minke whales in the North Pacific at next year's meeting.

North Pacific Bryde's whales

The Committee is in the process of developing initial *Implementation Simulation Trials* for western North Pacific Bryde's whales. In particular, it began a review of the reliability of catch statistics. Intersessional work will continue on this issue so that any uncertainty about such statistics can be incorporated into future trial structure. A full discussion of population structure and abundance-related issues will take place at next year's meeting and the Committee will determine whether the *pre-Implementation* stage of the process has been completed.

North Atlantic common minke whales

The Committee had expected to be in a position to undertake an *Implementation Review* of North Atlantic minke whales at the 2002 meeting. However, due to problems in making all of the data available suitably in advance of the meeting, it was agreed to postpone the review until next year. The review will primarily consider new information on stock structure and abundance.

Bycatches of large whales

The RMP estimates a limit for the number of non-natural removals, not simply a catch limit for commercial whaling. It is therefore important to estimate the numbers of whales removed from the population by indirect means including bycatches in fishing gear and ship strikes, for example.

The Scientific Committee began to consider this issue in some detail last year. It agreed that priority should be given to those areas where the RMP is likely to be implemented – such as the northwestern Pacific and the northeastern Atlantic. Four steps are required: (1) identification of the relevant fisheries; (2) description and categorisation of those fisheries to allow a sampling scheme to be devised; (3) identification of a suitable sampling strategy or strategies; and (4) design and implementation of the sampling scheme to enable estimation of the total bycatch.

The Committee has reviewed general methods for estimating bycatches. These fall under two headings: (1) those based on fisheries data and observer programmes; and (2) those based on genetic data. The former have been used successfully for several small cetacean populations. The Committee agreed that independent observer schemes are generally the most reliable means of estimating bycatch rates in a statistically rigorous manner, but that they may not always be practical and will require careful design.

The latter potentially represents a new way of estimating bycatches. The Committee has agreed that although genetic methods based on market samples may not be the primary approach to estimating bycatch, they could provide useful supplementary data that could not be obtained in another

way. The use of market samples to provide absolute estimates should not be ruled out. However, it will require further developments in sampling design with input from experts with detailed knowledge of market sampling issues. The possibility of holding a workshop on that subject is being considered.

This year, the Committee looked at the bycatches of large whales reported in National Progress Reports (progress reports are available on the IWC website: *http://www.iwcoffice.org/scweb/scprogreps*). Common minke whales were the most frequently reported species (> 230) with most records for Japan and eastern Korea. Compulsory reporting schemes exist in both these countries (it was voluntary in Japan prior to 1 July 2001). Possible reasons for the clumping of catches in these two areas was discussed but no clear explanation emerged, although lack of reporting by some countries is probably part of the explanation.

Work to further explore improved bycatch estimation methods for the two approaches noted above is continuing.

A major topic at this year's meeting concerned consideration of ways in which bycatches of large whales (and mortality of entangled whales) can be minimised. A report on this topic will be published in the 2003 *Supplement*.

DEVELOPMENT OF AN ABORIGINAL WHALING MANAGEMENT PROCEDURE

With the completion of the RMP, the Commission asked the Scientific Committee to begin the process of developing a new procedure for the management of aboriginal subsistence whaling. Such a procedure must take into account the different management objectives for such whaling when compared to commercial whaling. This is an iterative and ongoing effort. The Commission will establish an Aboriginal Whaling Scheme that comprises the scientific and logistical (e.g. inspection/observation) aspects of the management of all aboriginal fisheries. Within this, the scientific component might comprise some general aspects common to all fisheries (e.g. guidelines and requirements for surveys and for data c.f. the RMP) and an overall AWMP within which there will be common components and case-specific components.

At the 2002 meeting, the Committee completed its work with respect to the Bering-Chukchi-Beaufort Seas stock of bowhead whales. It agreed a *Strike Limit Algorithm (SLA)* for bowhead whales and the scientific aspects of a Scheme; this was adopted by the Commission. It noted that should the Commission decide, it would be possible to apply the *Bowhead SLA* at the current meeting. Work will continue intersessionally on gray whales and the Committee hopes to be able to present a formal recommendation to the Commission for a *Strike Limit Algorithm* for gray whales at the next meeting. The situation for the Greenlandic fisheries for fin and minke whales is less promising. A considerable amount of research, especially concerning stock identity, is required and to this end, the Committee has developed a research programme in cooperation with Greenlandic scientists.

ASSESSMENT OF STOCKS SUBJECT TO ABORIGINAL SUBSISTENCE WHALING

Aboriginal subsistence whaling is permitted for Denmark (Greenland, fin and minke whales), the Russian Federation (Siberia, gray and bowhead whales), St Vincent and The Grenadines (Bequia, humpback whales) and the USA (Alaska, bowhead and gray whales). It is the responsibility of the Committee to provide scientific advice on safe catch limits for such stocks and until the AWMP is completed then the Committee provides advice on a more *ad hoc* basis, carrying out major reviews according to the needs of the Commission in terms of establishing catch limits and the availability of data. It also carries out brief annual reviews of each stock.

The present catch limits had been set up to the 2002 season and so at the 2002 meeting, the Committee had to provide management advice for all of the stocks considered. The Commission sets catch limits based on the scientific advice and a 'need' statement from the countries involved.

Eastern gray whales

The primary assessment carried out was for the eastern gray whale population (Issue 1 of the present volume of the *Journal* was devoted to gray whale papers). New information on abundance, distribution, catches and ecology was presented. The population is believed to be close to carrying capacity. The Committee agreed that an annual take of up to 463 whales was acceptable; based on the submitted need statement, the Commission set a total for the 2003-6 seasons of 620 with a maximum of 140 in any one year.

Bering-Chukchi-Beaufort Seas stock of bowhead whales

In addition to the work on the *Bowhead SLA*, the Committee also examined the status of the Bering-Chukchi-Beaufort Seas stock of bowhead whales. New information included a preliminary abundance estimate for 2001 of 9,860 (95%CI 7,700 – 12,600) giving a rate of increase between 1978 and 2002 of 3.3% (95%CI 2%, 4.7%). The Committee noted that irrespective of its work on the bowhead *SLA*, the information here suggests that it is very likely that an annual catch of 102 whales will allow the stock to increase. Despite this, a proposal to continue to include provision for such catches (up to 280 bowhead whales to be landed in the period 2003 −2006, with no more than 68 whales struck in any year) failed to reach the necessary three-quarters majority in the Commission. The reason given by some of the 11 countries that voted no was that they believed Japan should also be allocated subsistence whales for four coastal whaling villages. They stressed that they also believed that the peoples of Alaska and Chukotka should also be granted their catch limits.

Minke and fin whales off West Greenland

The Committee received little new information on stocks of minke and fin whales off West Greenland this year. It has never been able to provide satisfactory management advice for these stocks and once again expressed great concern at this state of affairs. It stressed that obtaining adequate information on stock identity and abundance should be seen as of extremely high priority and made a number of research recommendations. Without this information, the Committee will not be able to provide safe management advice in accord with the Commission's management objectives, or develop a reliable *SLA* for many years, with potentially serious consequences for the status of the stocks involved. At the Commission, the same catch limits as previously in force were agreed for the 2003-6 period i.e. West Greenland minke whales – an annual limit of up to 175 strikes; East Greenland minke whales – an annual catch of up to 12 animals; West Greenland fin whales – an annual catch of up to 19 whales.

Humpback whales off St Vincent and the Grenadines
The Committee has been working on an in-depth assessment for North Atlantic humpback whales (see below). Based on the available data, the Committee believes it is most plausible that eastern Caribbean humpbacks are part of the West Indies breeding population (abundance in 1992/93 – 11,570, 95%CI 10,100 – 13,200). However, it recommended further collection of relevant data to confirm this. The Committee agreed that an annual catch of up to four whales was acceptable. After considerable debate in the Commission, a catch of up to 20 whales for the period 2003-7 was agreed (the Scientific Committee must review this in 2005).

STOCK IDENTITY

Of general concern to the assessment of any cetaceans is the question of stock identity and examination of this concept in the context of management plays an important role in much of the Committee's work, whether in the context of the RMP, AWMP or general conservation and management. In recognition of this, the Committee has established a Working Group to review theoretical and practical aspects of the stock concept in a management context. At the 2001 meeting, the Committee considered *inter alia:* terminology; stock structure in humpback whales; a range of analytical and statistical issues; the use of archetypes; and the combination of genetic and non-genetic information on stock identity.

This year, the Committee continued its work. In particular, it recognised the need to work towards an agreed definition of appropriate 'units-to-conserve' in a management context. Implicit in this is recognition that there may be need for case-by-case flexibility, and that it might be appropriate for the Committee to provide options and their implications when providing advice to the Commission. It is intended to have a full discussion of this idea next year. The Committee also examined a number of statistical and genetic issues relevant to this issue. Discussion focussed on use of 'traditional' hypothesis testing methods, a Bayesian approach (see Cui *et al.* in this volume) and a newer, as yet unpublished method (the boundary rank technique). In summary, the Committee noted that it is important, in any application of stock structure methods, to examine the sensitivity of conclusions to different *a priori* decisions about the definition of initial units, and about which population structure hypotheses to examine.

The Committee also recognised the importance of simulation testing to assess the performance of methods to identify population structure and will hold a specialist workshop to examine this in the coming year.

COMPREHENSIVE ASSESSMENT OF WHALE STOCKS

The 'Comprehensive Assessment' of whale stocks

The development of the concept of the 'Comprehensive Assessment' is reviewed in Donovan (1989). It can be considered as an in-depth evaluation of the status of all whale stocks in the light of management objectives and procedures; this would include the examination of current stock size, recent population trends, carrying capacity and productivity. Clearly, it is not possible to 'comprehensively assess' all whale stocks simultaneously, and the Committee has been working in an objective manner towards this, initially concentrating on stocks that have recently been or are presently subject to either commercial or aboriginal subsistence whaling. Some of these have already been discussed in the sections on the RMP and AWMP.

Antarctic minke whales
The Committee has carried out annual surveys in the Antarctic (south of 60°S) since the late 1970s. The last agreed estimates for each of the six management Areas for minke whales (see Donovan, 1991) were for the period 1982/83 to 1989/90 (IWC, 1991). At the 2000 meeting, the Committee agreed that whilst these represented the best estimates for the years surveyed, they were no longer appropriate as estimates of current abundance. An initial crude analysis of available recent data had suggested that current estimates might be appreciably lower than the previous estimates.

At the 2001 meeting, considerable time was spent considering Antarctic minke whales with a view to obtaining final estimates of abundance and considering any trend in these. This included a review of data sources and analytical methodology. After considering many of the factors affecting abundance estimates, there is still evidence of a decline in the abundance *estimates*, although it is not clear how this reflects any *actual* change in minke abundance. Three hypotheses that might explain these results were identified:

(1) a real change in minke abundance;
(2) changes in the proportion of the population present in the survey region at the time of the survey;
(3) changes in the survey process over time that compromise the comparability of estimates across years.

A considerable amount of work to investigate this further was undertaken at the 2002 meeting and a number of high priority tasks have been identified to be completed before the 2003 meeting.

Southern Hemisphere blue whales
The Committee is beginning the process of reviewing the status of Southern Hemisphere blue whales. An important part of this work is to try to develop methods to identify pygmy blue whales from 'true' blue whales at sea (IWC, 1999a) and progress is being made on this. Work on genetic and acoustic differentiation techniques is continuing and there is considerable progress with morphological methods. The Committee has agreed on a number of issues that need to be resolved before it is in a position to carry out an assessment, which it believes should commence in 2005.

Southern Hemisphere humpback whales
Considerable progress has been made in recent years in working towards an assessment of humpback whales. Attention has focussed both on data from historic whaling operations and on newly acquired photo-identification, biopsy and sightings data. The Committee made a number of research recommendations to further progress towards an assessment. An intersessional group has been established to review progress and determine whether it is feasible to set a deadline for the assessment to be completed.

North Atlantic humpback whales

At the 2001 meeting, priority was given to the Comprehensive Assessment of North Atlantic humpback whales. The Committee recognised the important contribution the international YoNAH (Years of the North Atlantic Humpback) project made to the assessment. This project combined photo-identification and molecular genetic techniques to collect as many photographs and skin biopsies as possible in four sampling periods over a wide geographical range during a period of two years (1992-1993). The principal objectives of the study were to increase understanding of: (a) abundance —both regionally and in total; (b) population genetic structure; (c) population spatial structure including rates of exchange among feeding grounds; and (d) reproductive behaviour and vital rates.

In reviewing population structure, the Committee concluded that North Atlantic humpback whales are characterised by relatively discrete feeding sub-stocks, with strong site fidelity by individuals. This latter factor also influences movement patterns within feeding grounds.

There is clear evidence for at least two breeding stocks in the North Atlantic. Whales from the western North Atlantic breed primarily in the West Indies, as do some whales that feed in the central North Atlantic. However, where other central North Atlantic animals and those from the Barents Sea breed is unknown.

The only breeding ground, other than the West Indies, known from historical and contemporary data is the Cape Verde Islands, but to date there is no direct evidence to support the idea that this is a breeding ground used by central and eastern North Atlantic animals. There may be a separate breeding population in the Norwegian Sea (as suggested in the late 1920s) and the possibility that there are three separate breeding stocks in the North Atlantic cannot be ruled out.

The Committee reviewed a number of population estimates for the feeding and breeding grounds.

This year, the Committee hoped to complete its assessment. It reviewed historical removals and agreed that the catch series was essentially complete for the 20th century although catches prior to then might be substantially underestimated. It also received new estimates of abundance from recent surveys in various parts of the North Atlantic. The Committee agreed that the abundance of the West Indies breeding population was around 10,800 in 1992/93 (see above) and was increasing at some 3% per year, at least between 1979 and 1992, the period for which suitable data are available. Attempts to model the population were unsuccessful (i.e. there was unacceptable model fit to the data) and a number of possible reasons for this were identified.

The Committee identified a number of research items that need to be completed before any further assessment is attempted.

North Atlantic right whales

The Committee has paid particular attention to the status of the North Atlantic right whale in the western North Atlantic in recent years (e.g. see Special Issue 2 of the Journal — *Right whales: worldwide status*). The Committee is extremely concerned about this population, which, whilst probably the only potentially viable population of this species, is in serious danger (*ca* 300 animals). By any management criteria applied by the IWC in terms of either commercial whaling or aboriginal subsistence whaling, there should be no direct anthropogenic removals from this stock.

This year, the Committee once again noted that individuals are continuing to die or become seriously injured as a result of becoming entangled in fishing gear or being struck by ships. It repeated that it is a matter of absolute urgency that every effort be made to reduce anthropogenic mortality in this population to zero. This is perhaps the only way in which its chances of survival can be directly improved. There is no need to wait for further research before implementing any currently available management actions that can reduce anthropogenic mortalities.

The Committee reviewed progress on a number of research and management recommendations concerning this stock.

Western North Pacific gray whales

This is one of the most endangered populations of great whales in the world. It numbers less than 100 animals (see the paper by Weller *et al.* in the last issue of the *Journal*, pp. 7-12) and there are a number of proposed oil and gas-related projects in and near its only known feeding ground. The Committee made a number of research and management recommendations for this population and will hold a Workshop in October to review this further. In conclusion, the Committee strongly reiterated that it is a matter of absolute urgency that every effort is made to reduce anthropogenic mortality (including direct catches) and disturbance to zero to save western North Pacific gray whales from extinction.

EFFECTS OF ENVIRONMENTAL CHANGE ON CETACEANS

There is an increasing awareness that whales should not be considered in isolation but as part of the marine environment; detrimental changes to their habitat may pose a serious threat to whale stocks. The Committee has examined this issue in the context of the RMP and agreed that the RMP adequately addresses such concerns. However, it has also emphasised that the species most vulnerable to environmental threats might well be those reduced to levels at which the RMP, even if applied, would result in zero catches (IWC, 1994). Over a period of several years, the Committee has developed two multi-national, multi-disciplinary research proposals. One of these, POLLUTION 2000+ (Reijnders *et al.*, 1999) has two aims: to determine whether predictive and quantitative relationships exist between biomarkers (of exposure to and/or effect of PCBs) and PCB levels in certain tissues; and to validate/calibrate sampling and analytical techniques. The other, SOWER 2000 (IWC, 2000) is to examine the influence of temporal and spatial variability in the physical and biological Antarctic environment on the distribution, abundance and migration of whales.

At the 2002 meeting, the Committee's primary topic concerned progress on the SOWER 2000 programme (IWC, 2000), particularly with respect to future collaboration with Southern Ocean GLOBEC and with CCAMLR. It also reviewed progress on the POLLUTION 2000+ programme (see Reijnders *et al.*, 1999). There was further discussion of the development of a report for the Commission that would provide an overview of regional environmental concerns and how best this might be achieved. A Workshop to address modelling-related issues related to the interactions between

cetaceans and fisheries was held in July 2002. The report of the Workshop will be published in next year's *Supplement*.

SMALL CETACEANS

Despite disagreement within the Commission over the management responsibilities of the IWC with respect to small cetaceans, it has been agreed that the Scientific Committee can study and provide advice on them. As part of this programme, the Committee has reviewed the biology and status of a number of species and carried out major reviews of significant directed and incidental catches of small cetaceans (Bjørge *et al*., 1994).

Last year, the Government of Japan had indicated that it would no longer co-operate with the Committee on small cetacean related matters. This year the Committee referred to the great value of the information provided by the Government of Japan on the status of small cetaceans in previous years and respectfully requested that the Government of Japan reconsider its position on this matter and resume the valuable contribution of Japanese scientists to its work on small cetaceans.

At the 2002 meeting, the Committee considered the status of humpback dolphins (genus *Sousa*). The taxonomy of the genus is somewhat confused, with up to five species being cited in various reports. Recognising the need for further taxonomic work, the Committee agreed to continue to recognise only two species at present: *S. teuszii*, the Atlantic humpback dolphin and *S. chinensis*, the Indo-Pacific humpback dolphin. Little information exists on the life history parameters of these essentially coastal species; that which does comes from South Africa and Hong Kong. Similarly, there is little information on abundance and trends. Actual and potential conservation problems are primarily due to habitat degradation and incidental capture in fishing and shark protection gear. Directed capture is relatively rare apart from Madagascar. The Committee concluded that there is insufficient information to assess the status of populations of this genus and it made a number of research recommendations.

The Committee also reviewed progress on previous recommendations it had made, particularly those concerning the critically endangered baiji and vaquita. Unfortunately, no new information was received on the baiji this year and the Committee has requested that information be provided next year. The Committee was informed of a new, integrated framework being developed to implement the recovery plan for the vaquita, and welcomed this new approach. It reiterated its endorsement of the primary conclusion of CIRVA (International Committee for the Recovery of the Vaquita) – that to ensure the future survival of the vaquita it will be necessary to eliminate all bycatches as rapidly as possible.

The Committee reviewed the draft report of the ASCOBANS recovery plan for harbour porpoises in the Baltic. It strongly endorsed the report and made some supplementary recommendations with respect to short-term pinger use.

The Committee also reviewed progress on the development of survey methodology for freshwater cetaceans and further work on the reduction of bycatches in fishing gear. No new information was received on the status of Dall's porpoises. Information on permits for takes of 1,000 white whales (for aboriginal subsistence purposes) and 10 killer whales (live-capture) by the Russian Federation were received. The Committee urged that assessment of the impact of such takes should be undertaken before their enactment.

Finally, the Committee repeated previous requests for all Governments to submit relevant information on direct and incidental catches of small cetaceans in their national progress reports.

SCIENTIFIC ASPECTS OF WHALEWATCHING

In 2000, the Committee had identified a number of areas for further research on possible long-term effects of whalewatching on whales and a number of possible data types that could be collected from whalewatching operations to assist in assessing their impact. The Committee developed this further at the 2002 meeting and will continue to work on data collection issues in the intersessional period.

The Committee also reviewed: information on noise from whalewatching vessels and aircraft, and any potential effects this might have on cetaceans; whalewatching guidelines and regulations; new information on dolphin feeding and 'swim-with' programmes.

REVIEW AND COMMENT ON SCIENTIFIC PERMITS ISSUED FOR SCIENTIFIC RESEARCH

All proposed scientific permits have to be submitted for review by the Scientific Committee following guidelines issued by the Commission. However, in accordance with the Convention the ultimate responsibility for issuing them lies with the member nation.

Most of the discussion at the 2002 meeting centred on reviewing the results of the two-year JARPN II feasibility study and the proposal for a further permit that involves taking 150 common minke whales, 50 Bryde's whales, 50 sei whales and 10 sperm whales each year for an unspecified period. The stated goal was to obtain information to contribute to the conservation and sustainable use of marine living resources in the western North Pacific. It includes sub-projects on: feeding ecology and ecosystems; monitoring of environmental pollutants in cetaceans and the marine ecosystem; further elucidation of stock structure. As for the feasibility study, there was considerable disagreement within the Committee over most aspects of this research programme, including objectives, methodology, sample sizes, likelihood of success, effect on stocks and the amount and quality of data that could be obtained using non-lethal research techniques.

The Committee also briefly considered the continuing programme on Antarctic minke whales that was last extensively reviewed in 1997 (IWC, 1998).

WHALE SANCTUARIES

The Committee had been asked by the Commission to review the Indian Ocean Sanctuary (IOS) and an intersessional working group had developed a proposed framework to carry out the review, in the light of guidelines developed by the Commission last year. The Committee's discussions of sanctuaries in the past have been somewhat inconclusive, with attention being drawn to a number of general arguments both in favour of and against sanctuary proposals. The discussion of the IOS had inevitably been coloured by these overall philosophical views. The Committee noted that lack of consensus in evaluating the scientific aspects of the IOS was not surprising considering that the sanctuary's original proposal did not clearly state its

scientific objectives. It recognised that the review process would benefit from explicitly stated objectives in Sanctuary proposals. However, while there was little consensus in evaluating the IOS, a considerable amount of substantive advice was provided on a number of sanctuary-related scientific issues.

The Committee considered a number of ways to improve the overall review process and priority will be given to this next year.

G.P. Donovan
Editor

REFERENCES

Bjørge, A., Brownell, R.L.J., Donovan, G.P. and Perrin, W.F. 1994. Significant direct and incidental catches of small cetaceans. A report by the Scientific Committee of the International Whaling Commission to the United Nations Conference on Environment and Development (UNCED). Appendix 2. List of small cetacean species. *Rep. int. Whal. Commn* (special issue) 15:127.

Donovan, G.P. 1989. Preface [to the volume *The Comprehensive Assessment of Whale Stocks*]. *Rep. int. Whal. Commn* (special issue) 11: iii-v.

Donovan, G.P. 1991. A review of IWC stock boundaries. *Rep. int. Whal. Commn* (special issue) 13:39-68.

Donovan, G.P. 2002. Editorial. *J. Cetacean Res. Manage.* 3(2): iii-vi.

International Whaling Commission. 1991. Report of the Scientific Committee. *Rep. int. Whal. Commn* 41:51-89.

International Whaling Commission. 1993. Chairman's Report of the Forty-Fourth Meeting, Appendix 3. Resolution on the Revised Management Scheme. *Rep. int. Whal. Commn* 43:40.

International Whaling Commission. 1994. Report of the Scientific Committee, Annex D. Report of the sub-committee on management procedures. *Rep. int. Whal. Commn* 44:74-92.

International Whaling Commission. 1998. Report of the Intersessional Working Group to Review Data and Results from Special Permit Research on Minke Whales in the Antarctic, Tokyo, 12-16 May 1997. *Rep. int. Whal. Commn* 48:377-412.

International Whaling Commission. 1999a. Report of the Scientific Committee. *J. Cetacean Res. Manage. (Suppl.)* 1:1-52.

International Whaling Commission. 1999b. Report of the Scientific Committee. Annex N. The Revised Management Procedure (RMP) for Baleen Whales. *J. Cetacean Res. Manage. (Suppl.)* 1:251-8.

International Whaling Commission. 2000. Report of the SOWER 2000 Workshop. *J. Cetacean Res. Manage. (Suppl.)* 2:319-46.

Reijnders, P.J.H., Aguilar, A. and Donovan, G.P. (eds.). 1999. *Chemical Pollutants and Cetaceans (Special Issue) 1*. G.P. Donovan (series ed.) International Whaling Commission, Cambridge, UK. v-viii + 273pp.

Small-scale population structure of eastern North Pacific harbour porpoises (*Phocoena phocoena*) indicated by molecular genetic analyses

S.J. CHIVERS*, A.E. DIZON*, P.J. GEARIN+ and K.M. ROBERTSON*

Contact e-mail: susan.chivers@noaa.gov

ABSTRACT

Concern about the conservation and management of harbour porpoise (*Phocoena phocoena*) populations, which have experienced relatively large incidental fishery kills in localised areas throughout their range, has prompted research to better understand their population structure. Both mitochondrial and nuclear (microsatellites) DNA were used to examine the intra-specific structure of harbour porpoise inhabiting the eastern North Pacific Ocean. Null hypotheses of panmixia were tested after mitochondrial DNA (mtDNA) control region sequence data (402 base pairs; $n = 249$) and allele frequency data (9 polymorphic loci; $n = 194$) were sub-divided into geographic strata defined *a priori*. Strata were based on sampling discontinuities and not discontinuities in population distribution. The mtDNA and nuclear gene data revealed statistically significant genetic differentiation between most strata ($\alpha = 0.05$) suggesting demographic independence of fairly small sub-units within the population. Since harbour porpoises are essentially continuously distributed in the eastern North Pacific, this degree of genetic differentiation was unexpected and needs to be considered in developing a sound management plan to protect them.

KEYWORDS: HARBOUR PORPOISE; CONSERVATION; GENETICS; NORTH PACIFIC; DISTRIBUTION; STOCK IDENTITY

INTRODUCTION

Concern over whether harbour porpoise (*Phocoena phocoena*) populations can sustain the level of observed incidental fisheries mortality has prompted research on their intra-specific population structure. Knowledge of this structure is necessary to properly assess the impact of fisheries mortalities by defining the appropriate regions for which to estimate animal abundance and incidental mortality. Recently published molecular genetic studies of harbour porpoises inhabiting the northwest Atlantic (Rosel *et al.*, 1999), the North and Baltic Seas (Tiedemann *et al.*, 1996; Wang and Berggren, 1997), West Greenland (Andersen *et al.*, 1997) and the seas around the UK (Walton, 1997) detected significant genetic differentiation between sampled strata. The detection of genetic differences between the geographically adjacent sampled areas in each study is quite remarkable, because harbour porpoises appear to be essentially continuously distributed and there are no apparent barriers to movement throughout the habitats they occupy.

Throughout their range, harbour porpoises are vulnerable to coastal gillnet fishing and the eastern North Pacific Ocean population focused on in this study is no exception (Gaskin, 1984; Perrin *et al.*, 1994). Harbour porpoises are locally abundant throughout their distribution, which in the eastern North Pacific extends from Point Conception, California around the North Pacific rim to the northern islands of Japan and as far north as Barrow, Alaska (Leatherwood *et al.*, 1983; Jefferson *et al.*, 1993). Their mortality has been well documented in coastal gillnet fisheries for halibut in California (Hanan *et al.*, 1993; Julian and Beeson, 1998) and in gillnet fisheries for salmon that operate near Spike Rock, Washington and in Puget Sound, Washington (Gearin *et al.*, 1994; 2000; Pierce *et al.*, 1996). However, the incidental take of harbour porpoises in gillnet fisheries operating around Vancouver Island, British Columbia and throughout Alaska is not well known because observer coverage is either lacking or less than 5% of the fishery (Barlow *et al.*, 1995a; Small and DeMaster, 1995). The common characteristic of these gillnet fisheries is that the fishing effort is generally intensive and localised, and if they operate in areas where harbour porpoises occur, the incidental take may be large.

The existing national management plan for the harbour porpoise in the eastern North Pacific Ocean recognises seven management units or stocks: (1) central California; (2) northern California; (3) Oregon/Washington coastal; (4) Washington inland waterways; (5) Southeast Alaska; (6) Gulf of Alaska; and (7) Bering Sea (Barlow *et al.*, 1995a; 1997; 1998; Hill *et al.*, 1996; Hill and DeMaster, 1998; Forney *et al.*, 1999). These management units cover fairly large geographic areas, and because fisheries mortality has been locally high within several of the units, there is concern about whether their scale is biologically appropriate and whether the boundaries are in the right place. Originally, these stocks were designated based on knowledge of the distribution of animals as well as analyses of mitochondrial DNA (mtDNA) control region sequences (Rosel *et al.*, 1995) and contaminant concentrations (Calambokidis and Barlow, 1991). Both of the studies concluded that there was evidence of limited dispersal between the strata represented in their study but that there was probably additional, finer structure within the populations (i.e. areas not yet sampled but inhabited by harbour porpoise). Management goals specified in the US Marine Mammal Protection Act of 1972 and its subsequent amendments for marine mammal populations occupying US territorial waters include, among others, maintaining populations as functional elements of their ecosystem. This has been interpreted to mean that a species' historical distribution and range should be maintained

* *NMFS, Southwest Fisheries Science Center, 8604 La Jolla Shores Dr., La Jolla, CA 92037, USA.*
+ *NMFS, National Marine Mammal Laboratory, 7600 Sand Point Way NE, Seattle, WA 98115, USA.*

(Barlow *et al.*, 1995b). Due to the limited understanding of population structure, whether the current management plan for the eastern North Pacific harbour porpoise population is adequate to meet these goals is uncertain.

This paper extends the existing knowledge by using molecular genetic techniques on a dataset that includes samples collected from additional areas that were not represented in previous studies. Both mitochondrial and nuclear gene markers were used to measure genetic differentiation between strata defined *a priori*. These markers quantify different aspects of gene flow, because they evolve at different rates and the mtDNA marker is maternally inherited, whereas, the nuclear markers are bi-parentally inherited. The maternal mode of inheritance for mtDNA means that the effective population size is approximately a quarter that of nuclear markers, which will result in more rapid differentiation of population sub-units, primarily due to genetic drift, when gene flow is limited (i.e. negligible movement of breeding females). Furthermore, the evolutionary rate of mtDNA makes it useful for reconstructing phylogeographic relationships, and these relationships were examined before performing intra-specific structure analyses to see whether an evolutionary process played an overall role in the population's structure. Analyses of both mitochondrial and nuclear DNA markers reveal patterns of gene flow, which can be used to infer movement and dispersal patterns of the breeding portion of the population studied, and thus provide evidence of intra-specific structure.

MATERIALS AND METHODS

Samples

Samples used in this study were collected along the west coast of the USA and Canada between 1984 and 1998 from animals incidentally taken in fisheries, found stranded on the beach or biopsied at sea (Fig. 1). The samples collected were predominantly skin tissue (92% skin; 8% muscle or internal organ tissue) preserved in a 20% dimethylsulphoxide solution saturated with NaCl (Amos and Hoelzel, 1991; Amos, 1997). All samples are stored in the Southwest Fisheries Science Center's Genetic Tissue Archive (contact author SJC for information).

DNA extraction

Standard molecular protocols were used to extract genomic DNA (Saiki *et al.*, 1988; Palumbi *et al.*, 1991). Extractions of DNA with a CTAB (cetyltrimethylammonium bromide) protocol (Winnepenninckx *et al.*, 1993) were successful for most samples, but when DNA yield was initially low, a phenol-chloroform technique was used for a second extraction (Sambrook *et al.*, 1989). Prior to amplification, the concentration of DNA extracted was determined spectrophotometrically and the purity assessed electrophoretically.

mtDNA amplification and sequencing

The 402 base pair region of the 5′ end of the hypervariable control region of the mtDNA gene was amplified using the polymerase chain reaction (PCR). The following primers were used: L15812 (5'-cctccctaagactcaagg-3') (Southwest Fisheries Science Center Laboratory, unpublished); or L15926 (5'acaccagtcttgtaaacc3'); and H16498 (5'cctgaagtaagaaccagatg3') (Rosel *et al.*, 1994), which are named according to their position in the mtDNA sequence of the fin whale (Árnason *et al.*, 1991). Standard protocols for the PCR were used with 50 µl reactions containing 1 µl (approximately 10-100ng) of genomic DNA, 37.75 µl MilliQ water, 5 µl of buffer (10mM Tris-HC1 (pH 8.3), 50mM KC1, 1.5mM $MgC1_2$), 3 µl of 10mM dNTP, 0.25 µl of *Taq* DNA polymerase, and 1.5 µl of 1 µM of each primer for the amplification. The PCR cycling was done on the *Perkin Elmer* 9600 thermocycler at 90°C for 2.5 min for the initial denaturation, followed by 35 cycles of 94°C for 45 sec, 48°C for 1 min and 72°C for 1.5 min, and a final extension at 72°C for 5 min (Saiki *et al.*, 1988). PCR products were cleaned using purification columns (QIAquick 250®;Qiagen).

Cycle sequencing was done using a profile of 25 cycles at 96°C for 10 sec, 50°C for 5 sec and 60°C for 4.0 min on the *Perkin Elmer* 9600 thermocycler. 12 µl sequencing reactions were used containing 3 µl of cleaned PCR product, 3 µl of primer (L15812 and H16498), 2 µl of PRISM® dRhodamine dye terminators (PE Applied Biosystems, Inc.) and 4µl MilliQ water. Both strands of the amplified DNA product of each specimen were sequenced independently as mutual controls using standard protocols on the Applied Biosystems Inc. (ABI) model 373 and 377 automated sequencers with most samples run on the ABI 377. All sequences were aligned by eye using SEQED, version 1.0.3 software (Applied Biosystems Inc., 1992).

Nuclear DNA processing

Nine dinucleotide primers were optimised for harbour porpoise: DlrFCB3, DlrFCB6 (Buchanan *et al.*, 1996), EV1, EV14, EV94, EV104 (Valsecchi and Amos, 1996), SL1026 (L. Garrison, Southwest Fisheries Science Center; pers. comm.), 415/416 and 417/418 (Andersen *et al.*, 1997). Extracted DNA was amplified using the PCR in 25 µl reactions containing 1 µl (approximately 10-100ng) genomic DNA, 18.25 µl water, 2.5 µl of buffer (same as sequencing buffer), 0.75 µl of each primer, 1.5 µl 10mM dNTP and 0.25 µl *Taq* DNA polymerase. All forward primers were labelled with a fluorescent dye. The thermal cycling profile for each locus was an initial 3 min at 97°C, followed by 35 amplification cycles of 30 sec at 90°C, 1 min at the specified annealing temperature for each primer (list follows) and 1 min at 72°C, and a final 5 min period at 72°C ensured extension of the PCR products. The optimal annealing temperature for each primer was 55°C for DlrFCB3, DlrFCB6, EV94 and SL1026, 49°C for EV1 and EV14, 48°C for EV104, and 45°C for 415/416 and 417/418. Size and purity of the amplicon was assessed electrophoretically. Successful amplifications were loaded onto an ABI 377 automatic sequencer for sizing with a commercial internal lane standard (ROX350®; PE Applied Biosystems Inc.). Allele fragment size was determined against a size standard using ABI's GENESCAN, version 3.1 software. The size of the allele is the number of repeat units × 2 plus the size of the flanking region for each base pair, and the size of each allelic pair for each loci constituted the raw data for analyses.

The primers listed above were selected from 22 that were optimised for use on harbour porpoises. For the 14 primers that were optimised, these were screened by plotting the sized alleles to ensure that dinucleotide repeats were amplified for all samples in the dataset and by testing each for the presence of non-amplifying loci, or so-called 'null' alleles. This additional screening was important because none of the primers were developed on the study species.

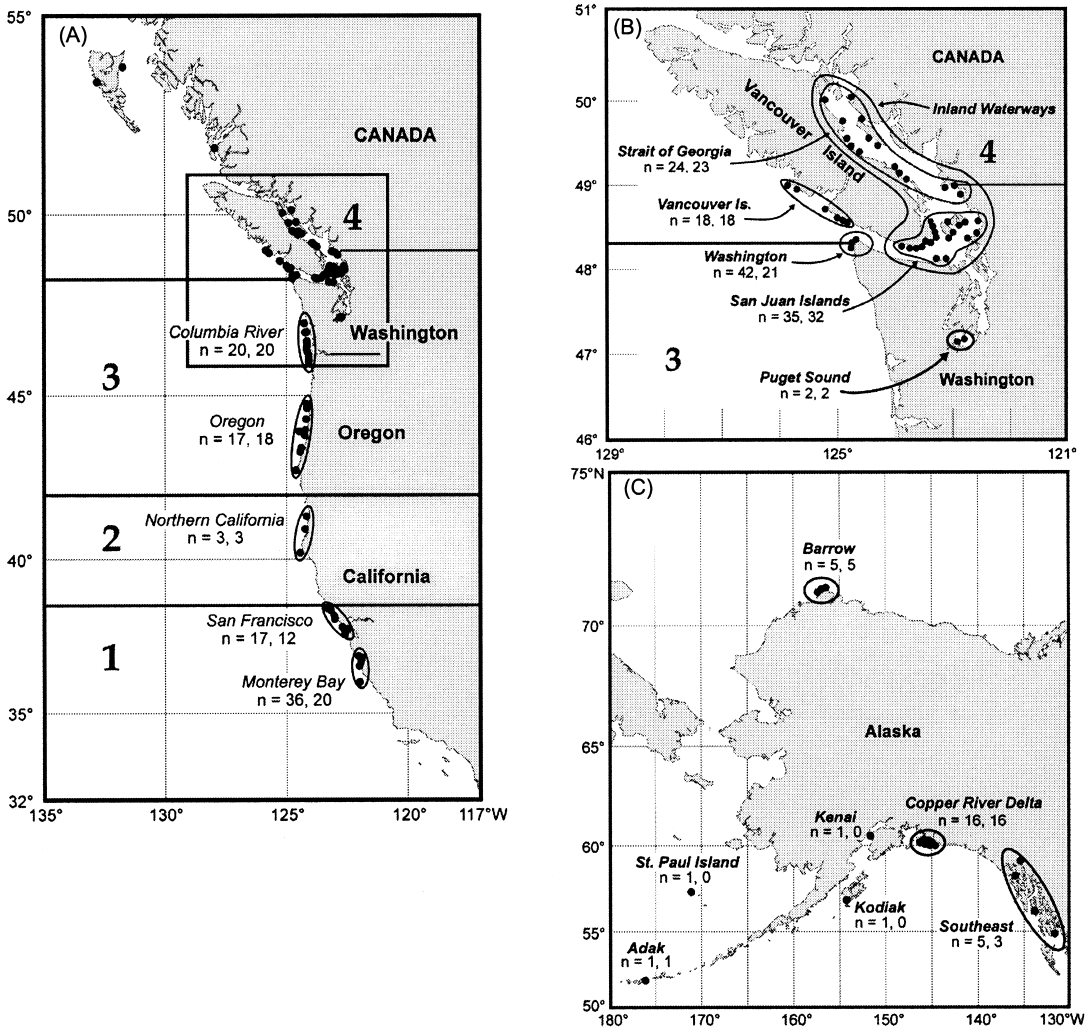


Fig. 1 Geographic locations of harbour porpoise specimens used in the molecular genetics study of population structure are plotted on three maps: (A) the California, Oregon and Washington coastal area; (B) the northern coast of Washington and the inland waterways of Washington and British Columbia, Canada; and (C) Alaska. The region outlined in (A) is expanded in (B) to show the detail of sample collection within that area. The large numbers and bold horizontal lines identify the existing management units and their boundaries: (1) central California; (2) northern California; (3) Oregon/Washington coastal; and (4) Washington inland waterways. Circles around groups of specimens identify each sampling stratum used in the genetic population structure analyses (see Methods for description). The name and sample size for each molecular marker (i.e. n = mtDNA sequence, nuclear DNA microsatellite) is listed next to each stratum. Discrepancies between the sample sizes of the mitochondrial and nuclear DNA datasets listed for a particular locale were due either to the inclusion of control region sequences for the Monterey Bay, California and northern Washington strata used in Rosel *et al.* (1995) for which microsatellite data were not available, or to low yield of DNA extracted from tissue samples of decomposed stranded animals that precluded sizing microsatellite loci.

Five primers were excluded from the dataset, because mutations in the alleles of one or more samples were detected. Null alleles were tested by adjusting annealing temperatures during PCR, because if null alleles are present, a heterozygotic state may appear as a homozygotic one under these conditions (Pemberton *et al.*, 1995; Jarne and Lagoda, 1996). The presence of null alleles was not detected in the nine primers selected for use in this study.

To further screen this dataset prior to analyses, individual relatedness was estimated using *Relatedness*, version 5.0.6 (Goodnight, 2000) and tested for evidence of linkage disequilibrium and deviations from Hardy-Weinberg equilibrium using *Arlequin*, version 2.0 (Schneider *et al.*, 2000). The estimated coefficients of relatedness (e.g. parent-offspring, full siblings) were calculated and used to identify duplicate individuals and first-order relatives in the dataset (Queller and Goodnight, 1989). Two parent-offspring pairs were identified and the offspring of each were removed from the dataset prior to analyses. No evidence of linkage disequilibrium was detected. However, in the tests of Hardy-Weinberg equilibrium, 2 of 72 comparisons (9 loci $\times$ 8 strata) were rejected. These results were not considered indicative of further sub-division within the strata in question, because the observed heterozygosity and allelic diversity was high for both, and furthermore, when making 72 separate tests, 5% would likely be rejected by chance alone.

Genetic variation analyses

For each sampling strata, the genetic variation of the control region was characterised by the number of unique haplotypes present and by estimates of haplotypic and nucleotide diversity (Nei and Tajima, 1981; Nei, 1987). Observed and expected heterozygosity for each locus processed was calculated using the procedure incorporated in Arlequin, version 2.0 (Schneider *et al.*, 2000).

Phylogeographic analyses

Prior to testing hypotheses of intra-specific structure, the control region sequences were examined for evidence of a phylogeographic (i.e. evolutionary) signal. In these analyses, no *a priori* assumptions about intra-specific structure are required, but inferences about phylogeographic patterns are made demonstrating concordance between related haplotypes and their sampling locations. The phylogeographic analyses included control region sequences from an additional 23 specimens from strata not used in the intra-specific structure analyses. For these strata, there were too few specimens to include with the other, more robustly sampled strata. The additional specimens included specimens collected off northern California ($n = 3$); Puget Sound, Washington ($n = 2$); northern British Columbia ($n = 4$); and several, geographically distant locations around Alaska ($n = 14$; Fig. 1). Only unique haplotypes were used and a minimum spanning network was generated with MINSPNET (Excoffier and Smouse, 1994) to construct the phylogeographic relationships. For reference, three haplotypes from Atlantic Ocean harbour porpoise and haplotypes for two sister species (*P. sinus* and *P. spinipinnis*) were included in the dataset. There were two dominant clades apparent in the optimal network, but there was no geographic concordance apparent in the network. Therefore, these results are not presented but are available from the authors.

Genetic population structure analyses

Conventional analyses designed to detect intra-specific structure are based on *a priori* stratification of the samples using non-genetic criteria (e.g. a distributional hiatus or geographic barriers). Therefore, the data were further analysed with *a priori* stratifications that sub-divided the dataset based on the current management scheme or sampling discontinuities. Different schemes were tested, because results from these analyses are fundamentally dependent on decisions about stratification.

The first *a priori* stratification of the dataset recognised the current management units: central California, Oregon/Washington coastal, Washington; and Washington inland waterways (Barlow *et al.*, 1995a; 1997; 1998). The existing northern California management unit was not represented in the analyses, since only three samples were collected from that area although it is known to have high densities of harbour porpoise (Forney, 1999).

The second *a priori* stratification was defined based on geographic sampling discontinuities and resulted in eight fairly fine-scale strata: Monterey Bay, California; San Francisco Bay and Russian River, California; central and southern Oregon; Columbia River, Oregon; northern Washington (Spike Rock); western shore of Vancouver Island, British Columbia; inland waterways of Washington and British Columbia; and Copper River Delta, Alaska (Fig. 1). These strata will be subsequently referred to as: Monterey Bay, San Francisco, Oregon, Columbia River, Washington, Vancouver Island, Inland waterways and Copper River Delta, respectively. The Inland waterways stratum covered a larger geographic range than any other strata in the study, and was therefore split into two: (1) the Strait of Georgia, Canada; and (2) the area south of the Strait of Georgia primarily around the San Juan Islands, Washington and southern tip of Vancouver Island. The null hypothesis of panmixia was tested for these two sub-strata, they were statistically distinguishable (see Results), and therefore, the second *a priori* stratification was modified to include sub-division of the Inland waterways. Thus, nine sampling strata were used in the fine-scale intra-specific analyses and the shorthand names used to reference the strata were: Monterey Bay, San Francisco, Oregon, Columbia River, Washington, Vancouver Island, San Juan Islands, Strait of Georgia and Copper River Delta. The Copper River Delta stratum was included as a 'reference' stratum (Fig. 1c). At the very least, genetic differences were expected between this geographically distant stratum and all others to the south.

Using the mtDNA data, the null hypothesis of panmixia was tested for intra-specific structure using both χ^2 and Φ_{ST}, because each statistic characterises a unique aspect of genetic differentiation. The χ^2 statistic detects differences in haplotype frequencies between strata and makes no assumptions about the evolution or relatedness of haplotypes (Rolf and Bentzen, 1989). On the other hand, Φ_{ST} detects differences in the relatedness of haplotypes between strata. That is, statistically significant Φ_{ST} values mean that haplotypes within a stratum are more closely related (i.e. have a smaller genetic distance or are more genetically homogenous) to each other than to those found in other strata. This statistic uses genetic distance to quantify relatedness and the number of homologous nucleotide differences between two individuals was used as the measure of genetic distance. Φ_{ST} is analogous to the more familiar F-statistic but is modified for pairwise comparisons of genetic distance data and tests significance with a non-parametric permutation method in an analysis of variance framework (AMOVA; Excoffier *et al.*, 1992).

Using the microsatellite data, the same *a priori* stratifications of data established for analyses of the mtDNA marker were tested. F_{ST} (Wright's fixation index; Wright, 1965; Cockerham and Weir, 1993) was the test statistic used in an AMOVA for the microsatellite data (Excoffier *et al.*, 1992). For analyses of both the mtDNA and microsatellite data, AMOVA was used as implemented in Arlequin, version 2.0 (Schneider *et al.*, 2000). For all intra-specific structure analyses, the null hypothesis was panmixia.

The Bonferonni multiple-test correction factor has become fairly routinely applied to results of genetic population structure analyses for cetacean species. However, the application of any correction factor needs to be carefully considered (see Rothman, 1990; Perneger, 1998; Bender and Lange, 1999). For example, one assumption of correction factors is that all null hypotheses are true simultaneously. When applied to genetic population structure analyses, the correction factor is routinely applied to all comparisons regardless of whether they are biologically plausible. In this study, the essentially linear coastal distribution of harbour porpoise means animals most likely move through adjacent (i.e. neighbouring) strata as they move along the coast, and therefore, only comparisons of adjacent strata would likely be biologically plausible, and the only comparisons that would likely be simultaneously true would be those with the Washington stratum, which has three neighbours: Columbia River, Vancouver Island and San Juan Islands. An additional consideration when applying correction factors is that they effectively reduce the critical value (α), or Type I error rate, but at the expense of the Type II error rate. In conservation management applications, reducing the Type I error rate means that one is more willing to commit an under-protection error (i.e. incorrectly pooling strata) than an over-protection error (i.e. incorrectly sub-dividing strata). Because the results of the analyses will likely be applied to management of the harbour porpoise, and the acceptance of

particular Type I and Type II error rates has implications for resource management (Dizon *et al.*, 1995; Taylor *et al.*, 1997), a correction factor was not applied to the analyses and the results were interpreted with $\alpha = 0.05$.

RESULTS

Genetic variation

Analysis of mtDNA data

There were 74 unique haplotypes identified with 88 variable sites among the 249 control region sequences in the dataset (Fig. 2). The overall nucleotide diversity was 0.014 and haplotypic diversity was 0.876. For each sampling strata, nucleotide diversity ranged from 0.0056-0.0243 and haplotypic diversity ranged from 0.377-0.956 (Table 1). Similarly, the smaller dataset used for the intra-specific structure analyses ($n = 225$: 115 males; 77 females; 33 unknown) had an estimated overall nucleotide diversity of 0.018 and haplotypic diversity of 0.876, and 63 of the 74 unique haplotypes identified in the complete dataset were represented (Table 2). As mentioned in the methods, 23 specimens were excluded from the intra-specific structure analyses, because there were too few specimens/locale to adequately represent separate strata that were geographically distant (Fig. 1).

Analysis of nuclear DNA data

All nine microsatellite loci used were polymorphic and the number of alleles per locus ranged from six for DlrFCB-6 to 24 for EV-1. There were 194 specimens in the dataset, of which 180 specimens (107 males; 73 females) were used in the analyses of intra-specific structure. The average observed heterozygosity was > 0.7 for each locus except 415/416 and 417/418, which had average observed heterozygosities of 0.689 and 0.509, respectively (Table 3).

Genetic population structure

Analysis of mtDNA data

For the first *a priori* stratification of data, which was the currently recognised management units: central California, Oregon/Washington coastal and Washington inland waterways, the overall Φ_{ST} was not statistically significant ($\Phi_{ST} = 0.014$, $P = 0.061$). Of the pairwise comparisons using Φ_{ST}, the null hypothesis was marginally rejected for only the central California versus Washington inland waterways comparison ($\Phi_{ST} = 0.034$, $P = 0.046$). However, statistically significant evidence was found of genetic differentiation in χ^2 for comparisons of the central California stratum with: (1) the Oregon/Washington coastal stratum (χ^2 $P < 0.001$); and with (2) Washington inland waterways stratum (χ^2 $P < 0.001$) (Table 4).

When the second *a priori* stratification of data was tested (i.e. Monterey Bay, San Francisco, Oregon, Columbia River, Washington, Vancouver Island, San Juan Islands, Strait of Georgia and Copper River Delta), more evidence of genetic differentiation was found. The overall Φ_{ST} was statistically significant ($\Phi_{ST} = 0.062$, $P < 0.0001$), and of the nearest neighbour comparisons considered most relevant to the question of intra-specific structure in this population, six of the nine were statistically significant using either χ^2 or Φ_{ST}. The comparisons of Columbia River to Washington, Washington to San Juan Islands and Vancouver Island to San Juan Islands were not statistically significant for either χ^2 or Φ_{ST} (Table 5).

Analysis of nuclear DNA data

When the current management units: central California, Oregon/Washington coastal and Washington inland waterways, were tested as population strata using the nuclear markers, the overall F_{ST} was not statistically significant ($F_{ST} = 0.0025$). However, the comparison of central California and Washington inland waterways strata was statistically significant ($F_{ST} = 0.0087$, $P = 0.020$) (Table 6).

When these data were analysed using our second *a priori* stratification (i.e. Monterey Bay, San Francisco, Oregon, Columbia River, Washington, Vancouver Island, San Juan Islands, Strait of Georgia and Copper River Delta), the overall F_{ST} was not statistically significant ($F_{ST} = 0.0075$). Furthermore, none of the nearest neighbour strata comparisons were statistically significant. It was expected that evidence of reproductive isolation for the Copper River Delta stratum would be found because it is geographically distant from all other strata. Significant genetic differences were detected in comparisons of the Copper River Delta to most strata, but it was not significantly distinguishable from the San Francisco, San Juan Islands and Strait of Georgia strata (Table 7). Similarly, significant differences were detected between comparisons of Monterey Bay and the strata most geographically distant to it. The Monterey Bay stratum is the one nearest the southern edge of the range of harbour porpoise in the eastern North Pacific, and therefore might be expected to be genetically distinguishable, if dispersal of breeding males and females is limited within the range.

DISCUSSION

The analyses of the mtDNA control region and nuclear DNA genetic markers of harbour porpoise provide evidence of a genetically sub-divided population organised into surprisingly small geographic units. Although results of analyses of the more broadly drawn first *a priori* stratification of data (i.e. the existing management units) provided evidence of intra-specific genetic distinctness, analyses of the final, most finely drawn *a priori* strata by and large demonstrated that each stratum was likely an isolated unit. Contrasting the results of these two analyses suggest that the current management units are likely composed of sub-units with unique genetic characteristics. In other words, the current management boundaries are drawn too broadly.

This conclusion was based on results from all of these analyses. The lack of statistical significance for the overall AMOVA results for both mtDNA and nuclear markers suggest that the within strata genetic variation is too great to be able to detect between strata differences, thus suggesting that structure exists within the current management units. The results from the second, more finely stratified dataset, were interpreted by considering detection of any significant difference in any genetic measurement between strata to be indicative of genetic differentiation (Table 8). Statistically significant genetic differentiation was detected with the mtDNA marker using either χ^2 or Φ_{ST} for six of our nine nearest neighbour comparisons. At least from a demographic perspective, these results indicate that there has been essentially no dispersal of breeding females between strata. The three comparisons that showed no evidence of genetic differentiation were Columbia River to Washington, Washington to San Juan Islands and San Juan Islands to Vancouver Island. These comparisons were all neighbours of each other, and possible explanations for the lack of evidence for genetic differentiation may be low statistical power or poorly defined strata that do not appropriately

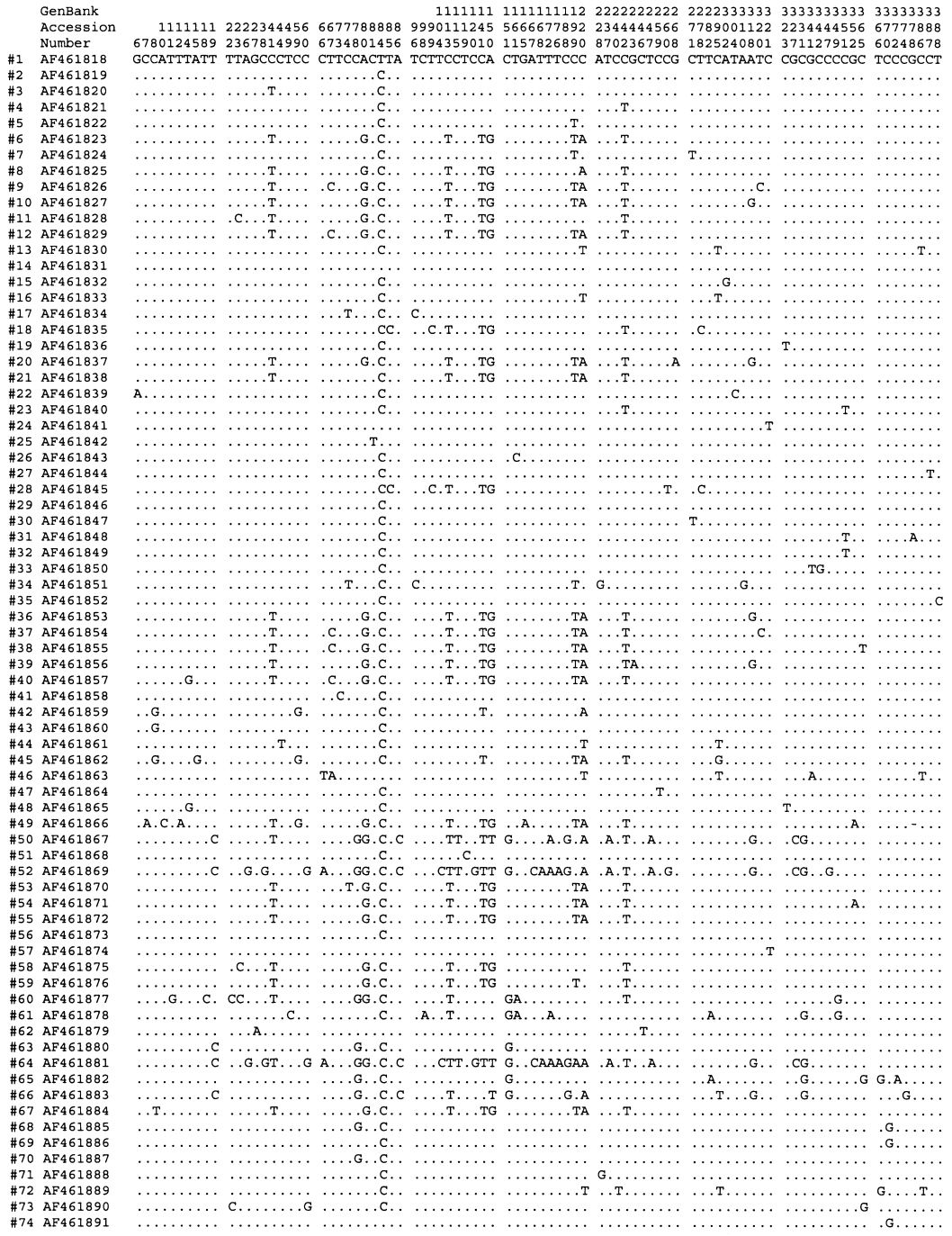


Fig. 2 The variable sites in 74 unique haplotypes identified for eastern North Pacific Ocean harbour porpoise.

Table 1

Sample size, number of haplotypes, haplotypic diversity and nucleotide diversity for each stratum.

Location	Sample size	No. of haplotypes	Haplotypic diversity	Nucleotide diversity
Monterey Bay, California	36	13	0.890	0.0153
San Francisco Bay, California	17	12	0.956	0.0162
Oregon (southern/central)	17	8	0.728	0.0083
Columbia River, Oregon	20	12	0.910	0.0243
Washington (Spike Rock)	42	14	0.892	0.0136
Inland waterways, Washington and British Columbia:				
San Juan Islands (and inland Strait of Juan de Fuca)	35	15	0.924	0.0186
Strait of Georgia, British Columbia	24	5	0.377	0.0061
Vancouver Island, British Columbia (western shore)	18	10	0.810	0.0127
Copper River Delta, Alaska	16	10	0.892	0.0056

Table 2
Frequencies for each unique mitochondrial DNA haplotype identified by sampling strata used in the intra-specific structure analyses.

Haplotype no.	Monterey Bay, CA	San Francisco, CA	Oregon (southern/central)	Columbia River, OR	Washington (Spike Rock)	Vancouver I., BC	San Juan Is, WA	Strait of Georgia, BC	Copper River Delta, AK	Frequency
1	0	0	0	0	0	0	0	0	1	1
2	8	3	9	5	10	8	7	19	5	74
3	0	0	0	0	0	0	0	0	3	3
4	0	0	0	0	3	0	0	0	1	4
5	0	0	0	2	1	0	2	0	1	6
6	5	2	0	1	6	0	3	0	0	17
7	0	2	0	0	0	0	0	0	0	2
8	1	1	0	0	3	0	0	0	0	5
9	1	2	0	0	0	0	0	0	0	3
10	0	1	0	0	1	0	1	0	0	3
11	0	1	1	4	2	1	4	1	0	14
12	4	0	0	0	0	0	0	0	0	4
13	3	0	1	0	3	2	4	0	0	13
14	7	0	0	0	0	0	0	0	0	7
15	1	1	1	1	0	0	0	0	0	4
16	0	0	0	0	0	0	2	1	0	3
17	0	0	0	1	2	1	1	0	0	5
18	0	0	0	1	2	0	0	0	0	3
19	0	0	1	0	1	0	0	0	0	2
20	0	0	0	0	0	0	0	2	0	2
21	0	0	2	0	0	0	1	0	0	3
22	0	0	0	0	0	0	2	0	0	2
23	0	0	0	0	0	0	4	0	0	4
24	0	0	0	0	2	0	0	0	0	2
25	1	0	0	0	0	0	0	0	0	1
26	0	0	0	0	0	0	0	1	0	1
27	1	0	0	0	0	0	0	0	0	1
28	1	0	0	0	0	0	0	0	0	1
29	0	0	0	0	0	0	0	0	1	1
30	0	0	0	0	0	0	0	0	1	1
31	0	0	0	0	0	0	0	0	1	1
32	0	0	0	0	0	0	0	0	0	0
33	0	0	0	1	0	0	0	0	0	1
34	0	0	1	0	0	0	0	0	0	1
35	0	0	0	0	0	0	0	0	0	0
36	0	0	1	0	0	0	0	0	0	1
37	0	0	0	0	0	0	0	0	0	0
38	1	0	0	0	0	0	0	0	0	1
39	1	0	0	0	0	0	0	0	0	1
40	1	0	0	0	0	0	0	0	0	1
41	0	0	0	0	0	0	0	0	0	0
42	0	0	0	0	0	0	0	0	0	0
43	0	0	0	0	0	0	0	0	1	1
44	0	0	0	0	0	0	0	0	0	0
45	0	0	0	1	0	0	0	0	0	1
46	0	0	0	0	0	0	0	0	0	0
47	0	0	0	0	0	0	1	0	0	1
48	0	0	0	0	0	0	1	0	0	1
49	0	0	0	0	0	0	0	0	0	0
50	0	0	0	0	0	1	0	0	0	1
51	0	0	0	0	0	1	0	0	0	1
52	0	0	0	0	0	1	0	0	0	1
53	0	0	0	1	0	0	0	0	0	1
54	0	0	0	1	0	0	0	0	0	1
55	0	0	0	0	0	0	1	0	0	1
56	0	0	0	0	0	0	1	0	0	1
57	0	0	0	1	0	0	0	0	0	1
58	0	1	0	0	0	0	0	0	0	1
59	0	0	0	0	0	0	0	0	0	0
60	0	0	0	0	0	0	0	0	0	0
61	0	0	0	0	0	0	0	0	1	1
62	0	1	0	0	0	0	0	0	0	1
63	0	1	0	0	0	0	0	0	0	1
64	0	1	0	0	0	0	0	0	0	1
65	0	0	0	0	0	1	0	0	0	1
66	0	0	0	0	0	1	0	0	0	1
67	0	0	0	0	0	1	0	0	0	1
68	0	1	0	0	0	0	0	0	0	0
69	0	0	0	0	1	0	0	0	0	1
70	0	0	0	0	1	0	0	0	0	1
71	0	0	0	0	1	0	0	0	0	1
72	0	0	0	0	1	0	0	0	0	1
73	0	0	0	0	1	0	0	0	0	1
74	0	0	0	0	1	0	0	0	0	1
Total	36	17	17	20	42	18	35	24	16	225

Table 3

Genetic diversity of the nuclear DNA dataset: the number of alleles, and the expected and observed heterozygosity of the microsatellite dataset for each locus by stratum.

Locus	No. alleles	Monterey Bay, CA (*n*=20)		San Francisco Bay, CA (*n*=12)		Oregon (south/central) (*n*=18)		Columbia River, OR (*n*=20)		Vancouver Island, BC (*n*=18)		Washington (Spike Rock) (*n*=21)		San Juan Islands, WA (*n*=32)		Strait of Georgia, BC (*n*=23)		Copper River Delta, AK (*n*=16)	
		H_E	H_O	H_E	H_O	H_E	H_O	H_E	H_O	H_E	H_O	H_E	H_O	H_E	H_O	H_E	H_O	H_E	H_O
DlrFCB-3	19	0.897	0.900	0.964	0.833	0.941	0.833	0.914	0.800	0.935	0.889	0.920	0.952	0.919	0.812	0.921	0.826	0.946	0.938
DlrFCB-6	6	0.788	0.850	0.862	0.917	0.776	0.833	0.796	0.800	0.752	0.667	0.816	0.762	0.772	0.844	0.751	0.783	0.758	0.812
EV-1	24	0.931	0.900	0.913	0.833	0.903	0.833	0.913	0.950	0.905	0.833	0.943	0.857	0.911	0.844	0.860	0.826	0.833	0.812
EV-14	8	0.622	0.600	0.750	0.833	0.611	0.667	0.833	0.950	0.763	0.833	0.700	0.714	0.743	0.688	0.735	0.739	0.744	0.750
EV-94	8	0.792	0.600	0.808	0.917	0.802	0.667	0.782	0.850	0.765	0.722	0.715	0.857	0.730	0.812	0.693	0.739	0.822	0.938
EV-104	13	0.873	0.850	0.793	0.917	0.849	0.722	0.883	1.000	0.859	0.889	0.908	0.857	0.905	0.906	0.908	0.870	0.879	0.812
SL10-26	16	0.865	0.750	0.840	0.833	0.790	0.889	0.838	0.900	0.914	0.889	0.915	0.857	0.891	0.938	0.857	0.913	0.897	0.812
415/416	7	0.740	0.850	0.627	0.667	0.697	0.778	0.660	0.600	0.808	0.611	0.653	0.714	0.653	0.812	0.638	0.609	0.605	0.562
417/418	7	0.654	0.550	0.475	0.417	0.527	0.500	0.533	0.450	0.522	0.500	0.645	0.428	0.496	0.531	0.501	0.391	0.613	0.812

Table 4

Results of intra-specific structure analyses using the currently recognised management units using the mitochondrial DNA control region sequences. In the upper diagonal are the Monte Carlo *P*-values for the comparison of test strata using χ^2, and in the lower diagonal are the Φ_{ST} test statistics with the corresponding *P*-value written in parentheses underneath. All text in cells with *P*-values $\leq$ 0.05 are printed in **bold** text and *P*-values between 0.05 and 0.10 are printed in *italics*.

	1. Central California (*n*=53)	2. Oregon/ Washington coastal (*n*=79)	3. Washington inland waterways (*n*=35)
1. Central California	-	**<0.001**	**<0.001**
2. Oregon/Washington coastal	*0.019 (0.063)*	-	**0.050**
3. Washington inland waterways	**0.034 (0.046)**	-0.002 (0.429)	-

reflect population distribution. Additional samples will be needed to resolve population structure in this region. Using the nuclear DNA markers, significant differences were detected only when the more distant strata were compared. These results are more difficult to interpret, because the effective population size is four times greater than that for mtDNA, but they suggest that breeding males move greater distances than breeding females within the study area. If this were true, even stronger evidence of genetic differentiation would be expected between strata using only females in the mtDNA dataset. However, the sample sizes for all *a priori* strata were too small to perform meaningful analyses using females only, because the dataset was approximately 60% male.

The rationale for combining evidence from the two markers: mtDNA and nuclear DNA, and from both statistics used to analyse the mtDNA data: χ^2 and Φ_{ST}, is that this approach is analogous to combining evidence from several disparate datasets, for example, morphological, contaminant and genetic studies. This type of approach has been applied in studies of intra-specific structure as a means to make an inference about animal movement patterns based on a preponderance of evidence (Dizon *et al.*, 1992). In such cases, you would not necessarily expect or demand each contrast to be significant for each criterion, but a significant finding in any marker provides relevant information. In this study, each genetic marker and statistic provides a different measure of genetic distinctness. The mtDNA marker provides information about the relative movement of breeding females, because it is maternally inherited, and the two statistics used provide information: (1) about the relative frequencies of haplotypes, which are expected to be different due to the complicated interplay of dispersal (albeit low) and genetic drift and would be detected by χ^2; and (2) about the relatedness, or evolution, of haplotypes, which change due to drift and mutation and would be detected by Φ_{ST} when there is essentially no gene flow between groups. On the other hand, nuclear markers are bi-parentally inherited, and therefore, provide information about the relative movements of both breeding males and females. When statistically significant differences are detected, they are interpreted as evidence that there is essentially no dispersal of males and females and that the strata are likely reproductively isolated.

The primary advantage of analysing molecular genetic data for evidence of population structure is that the data directly reflect gene flow and not transient, non-breeding interchange. The disadvantage is that the statistical power of the tests used to detect differences is inherently low (Dizon *et al.*, 1995; Taylor *et al*., 1997), and results of several studies (e.g. Slatkin, 1985; 1987; Slatkin and Barton, 1989; Hudson *et al.*, 1992) indicate that only near zero dispersal rates are reliably detected (i.e. approximately one migrant/generation). Therefore, when statistically significant differences in genetic variation are found between hypothesised population sub-units, the sub-units should be recognised as demographically distinct and managed separately. Even demographically insignificant amounts of dispersal will eliminate detectable genetic differences (Mills and Allendorf, 1996). While the interpretation of statistical significance is that there is effectively no movement of animals between strata compared (i.e. the dispersal rate was only a couple of animals per generation), the interpretation of negative results (i.e. no statistically significant differences) remains problematic. If an analysis fails to reject the null hypothesis of panmixia, the effective dispersal rate between population sub-units may still be several percent per year (e.g. 1-3%). At such low dispersal rates, population sub-units may be sufficiently isolated to warrant independent management, because the movement of animals between sub-units would be unlikely to compensate for anthropogenic mortality that exceeds the dispersal rate in at least one sub-unit (Taylor, 1997).

Estimating statistical power for genetic analyses is not straightforward and requires simulation modelling (Taylor *et al*., 1997). However, the role of statistical power in these

Table 5

Results of the fine-scale intra-specific structure analyses using Φ_{ST} and χ^2 for the mitochondrial DNA dataset. In the upper diagonal are the Monte Carlo *P*-values for the comparison of test strata using χ^2, and in the lower diagonal are the Φ_{ST} test statistics with the corresponding *P*-value written in parentheses underneath. All text in cells with *P*-values $\leq$ 0.05 are printed in **bold** text and *P*-values between 0.05 and 0.10 are printed in *italics*. Comparisons of nearest neighbours are on the off diagonal in bordered cells.

	1. Monterey Bay, CA	2. San Francisco, CA	3. Oregon (south/central)	4. Columbia River, OR	5. Washington (Spike Rock)	6. Vancouver Island, BC	7. San Juan Islands, WA	8. Strait of Georgia, BC	9. Copper River Delta, AK
	n=36	*n*=17	*n*=17	*n*=20	*n*=42	*n*=18	*n*=35	*n*=24	*n*=16
1. Monterey Bay, CA	-	**0.027**	**0.012**	**0.001**	**0.001**	**0.012**	**0.001**	**0.000**	**0.004**
2. San Francisco, CA	-0.016 (0.545)	-	*0.066*	0.258	*0.089*	*0.062*	**0.033**	**0.000**	*0.077*
3. Oregon (south\central)	**0.110 (0.023)**	**0.109 (0.036)**	-	0.158	0.118	0.806	*0.084*	*0.055*	*0.083*
4. Columbia River, OR	0.021 (0.154)	0.0006 (0.365)	**0.073 (0.043)**	-	0.164	0.176	0.259	**0.000**	0.129
5. Washington (Spike Rock)	0.005 (0.254)	-0.002 (0.372)	*0.067 (0.054)*	0.001 (0.334)	-	0.226	*0.055*	**0.000**	**0.048**
6. Vancouver I., BC	**0.132 (0.007)**	**0.126 (0.016)**	-0.012 (0.542)	**0.077 (0.018)**	**0.096 (0.013)**	-	0.144	**0.026**	*0.099*
7. San Juan Islands, WA	*0.033 (0.065)*	0.020 (0.156)	0.013 (0.224)	0.002 (0.352)	0.006 (0.249)	0.018 (0.158)	-	**0.001**	**0.012**
8. Strait of Georgia, BC	**0.157 (0.004)**	**0.170 (0.009)**	-0.027 (0.688)	**0.117 (0.006)**	**0.110 (0.010)**	-0.0002 (0.340)	**0.044 (0.049)**	-	**0.000**
9. Copper River Delta, AK	**0.189 (0.003)**	**0.211 (0.002)**	-0.002 (0.357)	**0.136 (0.002)**	**0.148 (0.003)**	0.019 (0.151)	**0.071 (0.016)**	-0.005 (0.442)	-

Table 6

Results for comparisons of the currently recognised management units using nuclear DNA. In the lower diagonal of this table are the F_{ST} test statistics with the corresponding *P*-value written in parentheses underneath. All text in cells with *P*-values $\leq$ 0.05 are printed in bold text.

	1. Central California (*n*=32)	2. Oregon/ Washington coastal (*n*=59)	3. Washington inland waterways (*n*=32)
1. Central California	-		
2. Oregon/Washington coastal	0.0012 (0.296)	-	
3. Washington inland waterways	**0.0087 (0.020)**	0.0003 (0.398)	-

analyses can be partially illustrated by comparing the number of statistically significant comparisons to sample size by stratum. The expectation is that the number of significant comparisons would increase with increasing sample size. In general, this is what is observed, but there are three notable exceptions: Monterey Bay, Strait of Georgia and Copper River Delta (Table 9). Both Monterey Bay and Copper River Delta were statistically distinguishable from all other strata in at least one analysis, even though Monterey Bay had the second largest mtDNA sample size and Copper River Delta had the smallest sample size (Table 8). In part, the genetic distinctness of these strata may be due to their location within the study area. Monterey Bay is at the southern extreme of the study area, while Copper River Delta is at the northern extreme. The Strait of Georgia stratum was statistically distinguishable in at least one analysis from all other strata except Oregon. The uniqueness of this stratum is likely due to the low haplotypic diversity observed (Table 1). Although a lack of statistical power due to relatively small sample sizes may help explain some of the inconsistencies in the results (i.e. not all comparisons were statistically significant when the nearest neighbour comparisons were significant), the discrepancies observed in the resolution of genetic population structure, particularly for the Oregon, Columbia River, Washington and San Juan Islands strata, likely have multiple explanations. For example, the comparisons of *a priori* strata, which were defined by sampling discontinuities, may not reflect the distribution of the population. Additional sampling in nearby areas together with a better understanding of regional distribution and seasonal movement patterns of the population will be needed to determine the influence of the *a priori* stratification in analyses. Additionally, even though χ^2 had been demonstrated to be a more powerful statistic than Φ_{ST} (Hudson *et al.*, 1992; Taylor and Chivers, 2000), this statistic did not appear to perform particularly well in the analyses (Table 9). Briefly, it is noted that sample size is not the only determinant of statistical power but that the number of unique haplotypes and observed haplotypic diversity also play a role. The performance of χ^2 observed most likely indicates that the sample sizes were likely too small for the relatively high haplotypic diversity/stratum (Table 1).

One of the goals of this study was to apply the results to management of the eastern North Pacific harbour porpoise population. However, achieving that goal remains elusive, because the samples were collected from discrete locales within the population's range, which do not necessarily correspond to the regional distribution of the population. In fact, some areas inhabited by large numbers of harbour porpoise were essentially un-sampled (e.g. northern California; Fig. 1a). The samples used were collected opportunistically in areas where people actively participate in stranding networks and where fishery observer programmes operate, and thus resulted in discrete geographic sampling. The discreteness of sampled areas ultimately dictated the *a priori* stratification of data for

Table 7

Results of the intra-specific population structure analyses using the nuclear DNA data. In the lower diagonal are the F_{ST} test statistics with the corresponding *P*-value written in parentheses underneath. All text in cells with *P*-values $\leq$ 0.05 are printed in bold text and *P*-values between 0.05 and 0.10 are printed in *italics*. Results for comparisons of nearest neighbor strata are in bordered cells.

	1. Monterey Bay, CA	2. San Francisco, CA	3. Oregon (south/central)	4. Columbia River, OR	5. Washington (Spike Rock)	6. Vancouver Island, BC	7. San Juan Islands, WA	8. Strait of Georgia, BC	9. Copper River Delta, AK
	n=20	*n*=12	*n*=18	*n*=20	*n*=21	*n*=18	*n*=32	*n*=23	*n*=16
1. Monterey Bay, CA	-								
2. San Francisco, CA	0.0030 (0.296)	-							
3. Oregon (south/central)	-0.0068 (0.884)	-0.0058 (0.761)	-						
4. Columbia River, OR	*0.0094* *(0.064)*	-0.0010 (0.504)	0.0001 (0.436)	-					
5. Washington (Spike Rock)	**0.0157** **(0.008)**	0.0072 (0.156)	**0.0142** **(0.019)**	*0.0084* *(0.075)*	-				
6. Vancouver I., BC	*0.0083* *(0.078)*	0.0077 (0.148)	-0.0004 (0.467)	-0.0010 (0.536)	*0.0081* *(0.084)*	-			
7. San Juan Islands, WA	**0.0108** **(0.021)**	0.0072 (0.138)	*0.0067* *(0.094)*	0.0002 (0.437)	0.0028 (0.243)	0.0055 (0.127)	-		
8. Strait of Georgia, BC	**0.0189** **(0.001)**	**0.0193** **(0.012)**	*0.0089* *(0.079)*	*0.0085* *(0.069)*	**0.0159** **(0.008)**	*0.0080* *(0.089)*	0.0040 (0.174)	-	
9. Copper River Delta, AK	**0.0111** **(0.048)**	*0.0126* *(0.070)*	**0.0140** **(0.026)**	**0.0129** **(0.037)**	**0.0107** **(0.050)**	**0.0180** **(0.005)**	0.0050 (0.181)	0.0059 (0.159)	-

Table 8

Combined results from all intra-specific structure analyses using mitochondrial (mtDNA) and nuclear DNA markers are presented in this table. The symbol for each test statistic: χ^2 and Φ_{ST} for mtDNA and F_{ST} for the nuclear DNA markers is printed in the cells for comparisons with $P \leq 0.05$. A "♦" denotes comparisons of sampling strata with no evidence of genetic differentiation. Nearest neighbour strata comparisons are the bordered cells.

	1. Monterey Bay, CA	2. San Francisco, CA	3. Oregon (south/central)	4. Columbia River, OR	5. Washington (Spike Rock)	6. Vancouver Island, BC	7. San Juan Islands, WA	8. Strait of Georgia, BC	9. Copper River Delta, AK
mtDNA	*n* = 36	*n* = 17	*n* = 17	*n* = 20	*n* = 42	*n* = 18	*n* = 35	*n* = 24	*n* = 16
Nuclear DNA	*n* = 20	*n* = 12	*n* = 18	*n* = 20	*n* = 21	*n* = 18	*n* = 32	*n* = 23	*n* = 16
1. Monterey Bay, CA	-								
2. San Francisco, CA	χ^2	-							
3. Oregon (south/central)	χ^2 Φ_{ST}	Φ_{ST}	-						
4. Columbia River, OR	χ^2	♦	Φ_{ST}	-					
5. Washington (Spike Rock)	χ^2 F_{ST}	♦	F_{ST}	♦	-				
6. Vancouver I., BC	χ^2 Φ_{ST}	Φ_{ST}	♦	Φ_{ST}	Φ_{ST}	-			
7. San Juan Islands, WA	χ^2 F_{ST}	χ^2	♦	♦	♦	♦	-		
8. Strait of Georgia, BC	χ^2 Φ_{ST} F_{ST}	χ^2 Φ_{ST} F_{ST}	♦	χ^2 Φ_{ST}	χ^2 Φ_{ST} F_{ST}	χ^2	χ^2 Φ_{ST}	-	
9. Copper River Delta, AK	χ^2 Φ_{ST} F_{ST}	Φ_{ST}	F_{ST}	Φ_{ST} F_{ST}	χ^2 Φ_{ST}	F_{ST}	χ^2 Φ_{ST}	χ^2	-

analyses, and thus, limits the conclusions regarding the true underlying intra-specific structure of the population and precludes precise placement of population sub-unit boundaries. However, dependence on opportunistic sampling means that it will be quite some time before there are sufficiently more samples to analyse and to refine existing knowledge.

Detecting genetic differences between neighbouring strata in these analyses was striking because harbour porpoises appear to be essentially continuously distributed in the

Table 9

Sample size and number of statistically significant ($P \leq 0.05$) comparisons detected by each statistic by stratum. The sampling strata are listed in order of decreasing sample size. The maximum number of significant comparisons possible was 24 or 8 per statistic.

Location	Sample size (mtDNA/ nuclear)	χ^2	Φ_{ST}	F_{ST}	Total no. significant comparisons
Washington (Spike Rock)	42/21	3	3	3	9
Monterey Bay, CA	36/20	8	4	4	16
Inland waterways, WA (Strait of Juan de Fuca and Juan Is)	35/32	4	2	1	7
Strait of Georgia, BC	24/23	7	5	3	15
Columbia River, OR	20/20	2	4	1	7
Vancouver I., BC (western shore)	18/18	2	4	1	7
San Francisco Bay, CA	17/12	3	4	1	8
Oregon (southern/central)	17/18	1	3	2	6
Copper River Delta, AK	16/16	4	5	4	13

eastern North Pacific Ocean and have high haplotypic diversity. Failure to find significant differences between all neighbouring strata in a single analysis should not be considered contradictory because the low statistical power of genetic analyses makes them sensitive to insufficient sample sizes. It is acknowledged that there are no obvious geographic barriers to restrict movement of animals in this population, and that knowledge about habitat preferences, movement patterns and seasonal distributions in the eastern North Pacific harbour porpoise is limited, but required, before sub-unit boundaries can be identified. However, the results of this study indicate that this population is likely highly stratified and that smaller management units would better preserve the population's intra-specific structure.

ACKNOWLEDGEMENTS

Specimens were made available to the Southwest Fisheries Science Center from a number of sources: J. Cordaro, Stranding Coordinator, Southwest Region, NMFS, Long Beach, CA; T. Price and D. Petersen, Observer Program, Southwest Region, NMFS, Long Beach, CA; J. Dines and P. Holahan, Humboldt State University, Arcata, CA; J. Hodder, Hatfield Biological Station, OR; D. Duffield, Portland State University, Portland, OR; B. Hanson and S. Osmek, National Marine Mammal Laboratory, Seattle, WA; J. Calambokidis, Cascadia Research Cooperative, Olympia, WA; R. Baird and T. Guenther, Marine Mammal Research Group, Victoria, British Columbia, Canada; K. Brix, Alaska Region, NMFS, Juneau, AK; R. Suydam, North Slope Borough, Barrow, AK; and V. Thayer, Southeast Fisheries Science Center, NMFS, Beaufort, NC. We thank all of them as well as many other participants whose names we do not know but who participate in the Marine Mammal Stranding Network that operates along the west coast of the United States for their dedication to collecting samples from stranded and fishery caught animals and for making the specimen material available for this study. We also thank Patty Rosel for making control region sequences available from her work on these harbour porpoise completed during the early 1990s. We further extend our thanks to Becky Nachenberg and Bethany Sutton for optimising primers and processing samples in the molecular genetics laboratory at the Southwest Fisheries Science Center with the assistance of Nick Kellar, Cristi Lux and Aviva Rosenberg who ably managed the genetics laboratory and ran the ABI sequencer for maximum productivity. This manuscript was improved by comments from all National Marine Mammal Laboratory scientists who work on harbour porpoise, members of the Pacific Scientific Review Group as well as Rick LeDuc, Sarah Mesnick, Barb Taylor and Robin Waples. This work was supported in part by funds received from the Office of Protected Resources, NMFS.

REFERENCES

Amos, W. 1997. Marine mammal tissue sample collection and preservation for genetic analyses. pp. 107-13. *In:* A.E. Dizon, S.J. Chivers and W.F. Perrin (eds.) *Molecular Genetics of Marine Mammals*. The Society for Marine Mammalogy, Lawrence, KS. 388pp.

Amos, W. and Hoelzel, A.R. 1991. Long-term preservation of whale skin for DNA analysis. *Rep. int. Whal. Commn* (special issue) 13:99-103.

Andersen, L.W., Holm, L.E., Siegismund, H.R., Clausen, B., Kinze, C.C. and Loeschke, V. 1997. A combined DNA-microsatellite and isozyme analysis of the population structure of the harbour porpoise in Danish waters and West Greenland. *Heredity* 78:270-6.

Árnason, U., Gullberg, A. and Widegren, B. 1991. The complete nucleotide sequence of the mitochondrial DNA of the fin whale, *Balaenoptera physalus*. *J. Mol. Evol.* 33(6):556-68.

Barlow, J., Brownell, R.L., DeMaster, D.P., Forney, K.A., Lowry, M.S., Osmek, S., Ragen, T.J., Reeves, R.R. and Small, R.J. 1995a. US Pacific marine mammal stock assessments. *NOAA Tech. Memo. NMFS-SWFSC* 219 162pp.

Barlow, J., Swartz, S.L., Eagle, T.C. and Wade, P.R. 1995b. US marine mammal stock assessments: guidelines for preparation, background, and a summary of the 1995 assessments. *NOAA Tech. Memo. NMFS-OPR* 6 73pp.

Barlow, J., Forney, K.A., Hill, P.S., Brownell, R.L., Jr., Carretta, J.V., DeMaster, D.P., Julian, F., Lowry, M.S., Ragen, T.J. and Reeves, R.R. 1997. U.S. Pacific marine mammal stock assessments: 1996. *NOAA Tech. Memo. NMFS-SWFSC* 248 223pp.

Barlow, J., Hill, P.S., Forney, K.A. and DeMaster, D.P. 1998. US Pacific marine mammal stock assessments: 1998. *NOAA Tech. Memo. NMFS-SWFSC* 258 40pp.

Bender, R. and Lange, S. 1999. Multiple test procedures other than Bonferroni's deserve wider use. *Br. Med. J.* 318:600.

Buchanan, F.C., Friesen, M.K., Littlejohn, R.P. and Clayton, J.W. 1996. Microsatellites from the beluga whale *Delphinapterus leucas*. *Mol. Ecol.* 5:571-5.

Calambokidis, J. and Barlow, J. 1991. Chlorinated hydrocarbon concentrations and their use for describing population discreteness in harbor porpoises from Washington, Oregon, and California. pp. 101-10. *In:* J.E. Reynolds III and D.K. Odell (eds.) *Marine Mammal Strandings in the United States. NOAA Tech. Rep. NMFS 98*. 157pp.

Cockerham, C.C. and Weir, B.S. 1993. Estimation of gene flow from F-statistics. *Evolution* 47:855-63.

Dizon, A.E., Lockyer, C., Perrin, W.F., DeMaster, D.P. and Sisson, J. 1992. Rethinking the stock concept: a phylogeographic approach. *Conserv. Biol.* 6(1):24-36.

Dizon, A.E., Taylor, B.L. and O'Corry-Crowe, G.M. 1995. Why statistical power is necessary to link analyses of molecular variation to decisions about population structure. pp. 288-94. *In:* J.L. Nielsen (ed.) *Evolution and the Aquatic Ecosystem: Defining Unique Units in Population Conservation. American Fisheries Science Symposium 17*. American Fisheries Society, Bethseda, MD.

Excoffier, L. and Smouse, P.E. 1994. Using allele frequencies and geographic subdivision to reconstruct gene trees within a species: molecular variance parsimony. *Genetics* 136:343-59.

Excoffier, L., Smouse, P.E. and Quattro, J.M. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131:479-91.

Forney, K.A. 1999. The abundance of California harbor porpoise estimated from 1993-97 aerial line-transect surveys. *NOAA NMFS SWFSC Admin. Rep.* LJ-99-02. 16pp. [Available from Southwest Fisheries Science Center, 8604 La Jolla Shores Dr., La Jolla, CA 92037].

Forney, K.A., Muto, M.M. and Baker, J. 1999. US Pacific marine mammal stock assessment: 1999. *NOAA Tech. Memo. NMFS-SWFSC* 282 62pp.

Gaskin, D.E. 1984. The harbour porpoise, *Phocoena phocoena* (L.): regional populations, status and information on direct and indirect catches. *Rep. int. Whal. Commn* 34:569-86.

Gearin, P.J., Melin, S.R., DeLong, R.L., Kajimura, H. and Johnson, M.A. 1994. Harbor porpoise interactions with a Chinook salmon set-net fishery in Washington State. *Rep. int. Whal. Commn* (special issue) 15:427-38.

Gearin, P.J., Gosho, M.E., Laake, J., Cooke, L., Delong, R.L. and Hughes, K.M. 2000. Experimental testing of acoustic alarms (pingers) to reduce bycatch of harbour porpoise, *Phocoena phocoena*, in the state of Washington. *J. Cetacean Res. Manage.* 2(1):1-10.

Goodnight, K.F. 2000. Relatedness, ver. 2.000. [Available at *http://gsoft.smu.edu/GSoft.html*].

Hanan, D.A., Holts, D.B. and Coan, A.L. 1993. The California drift gillnet fishery for sharks and swordfish, 1981-82 through 1990-91. *California Dept. of Fish and Game Fish Bull.* 175:1-95.

Hill, P.S. and DeMaster, D.P. 1998. Alaska marine mammal stock assessments, 1998. *NOAA Tech. Mem. NMFS-AFSC*- 97 166pp.

Hill, P.S., DeMaster, D.P. and Small, R.J. 1996. Alaska Marine Mammal Stock Assessments. *NOAA Tech. Mem. NMFS-AFSC*- 97 150pp.

Hudson, R.R., Boos, D.D. and Kaplan, N.L. 1992. A statistical test for detecting geographic subdivision. *Mol. Biol. Evol.* 9:138-51.

Jarne, P. and Lagoda, P.J.L. 1996. Microsatellites, from molecules to populations and back. *Trends Ecol. Evol.* 11:424-9.

Jefferson, T.A., Leatherwood, S. and Webber, M.A. 1993. *FAO Species Identification Guide. Marine Mammals of the World*. FAO, Rome. 320pp.

Julian, F. and Beeson, M. 1998. Estimates of marine mammal, turtle, and seabird mortality for two California gillnet fisheries: 1990-1995. *Fish. Bull.* 96:271-84.

Leatherwood, S., Reeves, R.R. and Foster, L. 1983. *The Sierra Club Handbook of Whales and Dolphins*. Sierra Club Books, San Francisco. xvii+302pp.

Mills, L.S. and Allendorf, F.W. 1996. The one-migrant-per-generation rule in conservation and management. *Conserv. Biol.* 10:1509-18.

Nei, M. (ed.). 1987. *Molecular Evolutionary Genetics*. Columbia University Press, New York. x+512pp.

Nei, M. and Tajima, F. 1981. DNA polymorphism detectable by restriction endonucleases. *Genetics* 97:145-63.

Palumbi, S.R., Martin, A.P., Romero, S., McMillan, W.O., Stice, L. and Grawbowski, G. 1991. *The Simple Fool's Guide to PCR Version 2.0*. University of Hawaii, Honolulu. 44pp.

Pemberton, J.M., Slate, J., Bancroft, D.R. and Barrett, J.A. 1995. Nonamplifying alleles at microsatellite loci: a caution for parentage and population studies. *Mol. Ecol.* 4:249-52.

Perneger, T.V. 1998. What's wrong with Bonferroni adjustments. *Br. Med. J.* 316:1236-8.

Perrin, W.F., Donovan, G.P. and Barlow, J. (eds.). 1994. *Report of the International Whaling Commission (Special Issue 15). Gillnets and Cetaceans*. International Whaling Commission, Cambridge, UK. 629pp.

Pierce, J.D., Alexandersdottir, M., Jeffries, S.J., Erstad, P., Beattie, W. and Chapman, A. 1996. Interaction of marbled murrelets and marine mammals with the 1994 Puget Sound sockeye gill net fishery. Final report to Washington Department of Fish and Wildlife, Olympia, WA. 21pp.

Queller, D.C. and Goodnight, K.F. 1989. Estimating relatedness using genetic markers. *Evolution* 43:258-75.

Rolf, D.A. and Bentzen, P. 1989. The statistical analysis of mitochondrial DNA polymorphisms: X^2 and the problem of small samples. *Mol. Biol. Evol.* 6:539-45.

Rosel, P.E., Dizon, A.E. and Heyning, J.E. 1994. Genetic analysis of sympatric morphotypes of common dolphins (genus: Delphinus). *Mar. Biol.* 119(2):159-67.

Rosel, P.E., Dizon, A.E. and Haygood, M.G. 1995. Variability of the mitochondrial control region in populations of the harbour porpoises, *Phocoena phocoena*, on interoceanic and regional scales. *Can. J. Fish. Aquat. Sci.* 52:1210-9.

Rosel, P.E., France, S.C., Wang, J.Y. and Kocher, T.D. 1999. Genetic structure of harbour porpoise *Phocoena phocoena* populations in the Northwest Atlantic based on mitochondrial and nuclear markers. *Mol. Ecol.* 8:S41-54.

Rothman, K.J. 1990. No adjustments are needed for multiple comparisons. *Epidemiology* 1:43-6.

Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B. and Erlich, H.A. 1988. Primer-directed amplification of DNA with a thermostable DNA polymerase. *Science* 239:487-91.

Sambrook, J., Fritsch, E.F. and Maniatis, T. 1989. *Molecular Cloning: A Laboratory Manual*. 2nd Edn. Cold Spring Harbor Laboratory, New York.

Schneider, S., Roessli, D. and Excoffier, L. 2000. *Arlequin ver. 2.000: A Software for Population Genetics Data Analysis*. Genetics and Biometry Laboratory, University of Switzerland, Geneva. [Available at: *http://anthro.unige.ch/arlequin*].

Slatkin, M. 1985. Gene flow in natural populations. *Annu. Rev. Ecol. Syst.* 16:393-430.

Slatkin, M. 1987. Gene flow and the geographic structure of natural populations. *Science* 236:787-92.

Slatkin, M. and Barton, N.H. 1989. A comparison of three indirect methods for estimating average levels of gene flow. *Evolution* 43:1349-68.

Small, R.J. and DeMaster, D.P. 1995. Alaska marine mammal stock assessments 1995. *NOAA Tech. Mem. NMFS-AFSC*- 57:93pp.

Taylor, B.L. 1997. Defining 'population' to meet management objectives for marine mammals. pp. 49-65. *In:* A.E. Dizon, S.J. Chivers and W.F. Perrin (eds.) *Molecular Genetics of Marine Mammals*. The Society for Marine Mammalogy, Lawrence, KS. 388pp.

Taylor, B.L. and Chivers, S.J. 2000. Evaluating the performance of different statistics to detect population subdivision. Paper SC/F2K/J5 presented at the JARPN Review Meeting, Tokyo, Japan, 7-10 February 2000 (unpublished). 10pp. [Paper available from the Office of this Journal].

Taylor, B.L., Chivers, S.J. and Dizon, A.E. 1997. Using statistical power to interpret genetic data to define management units for marine mammals. pp. 347-64. *In:* A.E. Dizon, S.J. Chivers and W.F. Perrin (eds.) *Molecular Genetics of Marine Mammals*. The Society for Marine Mammalogy, Lawrence, KS. 388pp.

Tiedemann, R., Harder, J., Gmeiner, C. and Haase, E. 1996. Mitochondrial DNA sequence patterns of harbour porpoise (*Phocoena phocoena*) from the North and the Baltic Sea. *Z. Saugetierkd* 61:104-11.

Valsecchi, E. and Amos, W. 1996. Microsatellite markers for the study of cetacean populations. *Mol. Ecol.* 5:151-6.

Walton, M.J. 1997. Population structure of harbour porpoises *Phocoena phocoena* in the seas around the UK and adjacent waters. *Proc. R. Soc. Lond. Ser. B.* 264:89-94.

Wang, J.Y. and Berggren, P. 1997. Mitochondrial DNA analysis of harbour porpoises (*Phocoena phocoena*) in the Baltic Sea, the Kattegat-Skagerrak Seas and off the west coast of Norway. *Mar. Biol.* 127:531-7.

Winnepenninckx, B., Backeljau, T. and DeWachter, R. 1993. Extraction of high molecular weight DNA from mollusks. *Trends in Genetics* 9:407.

Wright, S. 1965. The interpretation of population structure by F-statistics with special regard to systems of mating. *Evolution* 19:395-420.

Bayes and Empirical Bayes approaches to addressing stock structure questions using mtDNA data, with an illustrative application to North Pacific minke whales

GURONG CUI*, ANDRÉ E. PUNT+*, LUIS A. PASTENE # AND MUTSUO GOTO#

Contact e-mail: Peter.Cui@csiro.au

ABSTRACT

Bayesian methods using mtDNA data are developed to compare single- and multiple-stock hypotheses. The likelihood of the data is assumed to be multinomial and the multivariate prior for the probability of an individual having a particular haplotype is assumed to be of the Dirichlet-β form. The values for the parameters of this prior are either determined using an Empirical Bayes approach or assumed to be distributed according to a log-normal hyper-prior (the 'Full Bayes' approach). The Empirical and Full Bayes methods are examined using simulation. The performance of the Empirical Bayes method is found to be much worse than that of the Full Bayes method. Illustrative comparisons for North Pacific minke whales based on the latter method confirm previous results that sub-areas 6 and 7 contain different stocks. Results of the application of this method to the mtDNA data for the sub-areas to the east of Japan, although generally uninformative, are nevertheless consistent with analyses based on hypothesis testing using allozymes and mtDNA. The results from this method should, however, be used for management purposes with some caution. This is because, although some testing of the Full Bayes method has been completed and suggests that when applied to data for two stocks that differ substantially in haplotype frequency, or when sample sizes are large and there is only one stock, performance is adequate, in common with most other methods for analysing genetics data, its performance has yet to be fully evaluated.

KEYWORDS: BAYES; GENETICS; STOCK IDENTITY; NORTHERN HEMISPHERE; NORTH PACIFIC OCEAN; MINKE WHALE

INTRODUCTION

One of the key uncertainties identified during the development of the Revised Management Procedure (RMP) for baleen whales was that of uncertainty regarding stock structure (IWC, 1992; Hall and Donovan, 2001). Fig. 1 illustrates the problem generically. Areas A and B are areas covered during abundance surveys while all of the historical catch is taken from Area A. The future intention of the fishery is to operate in Area A (which may, for example, be the closest to port). The catch limit for Area A can be based either on (1) the survey and catch data for Area A only, or (2) on the survey and catch data for both Areas combined.

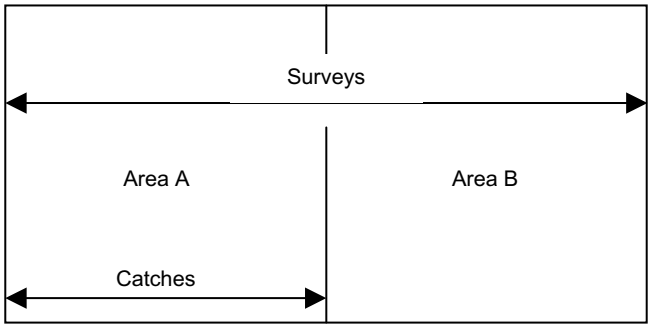


Fig. 1. The generic one-stock / two-stock problem. The historical (and future) catches are taken from Area A and surveys cover both Areas A and B. It is uncertain whether Areas A and B contain one or two stocks (and in the latter case in a manner that their relative proportions in Areas A and B differ appreciably).

Option (1) is appropriate if separate stocks[1] are found in Areas A and B, whilst option (2) is appropriate if there is only a single stock. If the catch limit is based on data for Area A only and there is in fact a single stock, the resource will be underutilised. Conversely, if it is based on the data for both Areas when Area A contains a separate stock from Area B, overexploitation will occur in Area A. The RMP can overcome some of the problems associated with uncertainty about stock structure through its catch capping and catch cascading options (IWC, 1994), but will yield improved performance (better catches for the same perceived risk) if some of this uncertainty can be resolved.

One of the most common recent approaches to attempt to resolve stock structure questions is collection and analysis of genetics data (e.g. see IWC, 1991; Dizon *et al.*, 1997). Traditionally, this has been examined using classical (frequentist) statistical methods based on the null hypothesis of panmixia. This classical approach is based on well-established statistical techniques (e.g. Excoffier *et al.*, 1992; Hudson *et al.*, 1992). Statistics related both to haplotypes (based on the haplotype frequencies only) and to sequencing (based on haplotype frequencies and genetic distances among haplotypes) are used. These techniques provide clear guidance regarding the most appropriate stock

[1] For simplicity, this paper discusses stock differentiation issues in the context of the existence of a stock boundary which exactly and completely separates stocks. In reality, of course, there will be a region of overlap, and any boundary line specified would constitute a trade-off choice which attempts to minimise the proportions of each stock likely to be present on each's 'other' side of that line.

* *CSIRO Division of Marine Research, GPO Box 1538, Hobart, TAS 7001, Australia.*
+ *School of Aquatic and Fishery Sciences, Box 355020, University of Washington, Seattle, WA 98195-5020, USA.*
# *Institute of Cetacean Research, Toyomi-cho 4-5, Chuo-ku, Tokyo 104-0055, Japan.*

structure hypothesis if a statistically significant result is obtained. However, the implications of a non-significant result are unclear. This is because a non-significant result can arise either because there is only a single stock in the area from which the data were collected, or because there is stock structure in the area but the sample size is too small to detect this. Furthermore, a non-significant result could also arise if there are really two stocks but the boundary between the strata chosen for data analysis does not correspond to that between the stocks.

In principle, the use of Bayesian techniques can overcome these problems. The outcome of a Bayesian comparison of single- and two-stock hypotheses (models) is the posterior odds ratio (Jeffreys, 1961; Kass and Raftery, 1995). The posterior odds ratio is the ratio of the relative probability of the one- to the two-stock hypothesis. Thus, a very large value will indicate preference for the one-stock hypothesis, a value close to zero preference for the two-stock hypothesis, and a value close to unity preference for neither hypothesis. The posterior odds ratio is the product of the prior odds ratio and the Bayes Factor. For the illustrative calculations of this paper, the prior odds ratio is assumed to be unity (i.e. the one- and two-stock hypotheses are equally likely *a priori*) so the posterior odds ratio is equal to the Bayes Factor.

Use of Bayesian methods is perhaps preferable to the use of classical statistical methods in any case. This is because they provide the (relative) probability of alternative hypotheses rather than simply the ability to reject one of the two models at some pre-specified level of type I error. Determining the relative probability of alternative hypotheses is preferable because it avoids the need for the specification of a somewhat arbitrary level of type I error, and because risk in fisheries management is related not only to the probability of an event but also to the severity of possible outcomes given that event. Thus, a stock structure hypothesis that has major management implications may warrant consideration by the decision makers even if it has relatively low probability.

Bayesian methods are being used increasingly to analyse genetics data (e.g. Lulhart and England, 1999; Shoemaker *et al.*, 1999; Kitada *et al.*, 2000; Pella and Masuda, 2001) and hence to determine the relative probabilities of alternative stock structure hypotheses. Punt *et al.* (2000) developed an approach for determining the relative probability of alternative stock structure hypotheses using allozyme data. Although allozyme data have been widely used in studies of stock structure (e.g. Butterworth *et al.*, 1996; Gardner and Ward, 1998), allozymes mutate at a slower rate than mtDNA and microsatellites so they have lower power to detect genetic differences (Bossart and Pashley Powell, 1998). This paper therefore develops a Bayesian framework within which single- and multiple-stock hypotheses can be compared using mtDNA data. The approach is evaluated using simulation and then, for illustrative purposes, applied to data for North Pacific minke whales (see Fig. 2 for the management sub-areas defined for North Pacific minke whales).

METHODS

Basic formulation

The region to be sampled (and for which stock structure hypotheses are postulated) is assumed to be divided into n sub-areas. For each sub-area i, a number, N^i, of animals are sampled. This leads to a total dataset $\{x_j^i: j = 1, 2, \ldots, k; i = 1, 2, \ldots, n\}$ where x_j^i is the number of animals sampled in sub-area i that have haplotype j, and k is the total number of haplotypes in the whole dataset (by definition $\sum_{j=1}^{k} x_j^i = N^i$).

Given a random sampling scheme (as is the case for the mtDNA data for North Pacific minke whales; e.g. Fujise (2000)), the dataset for sub-area i can be considered to be a multinomial sample from the population in sub-area i. If p_j^i is the proportion of animals in sub-area i with haplotype j, then $\underline{x}^i \sim MN(\underline{p}^i, N^i)$ and the likelihood for the dataset for

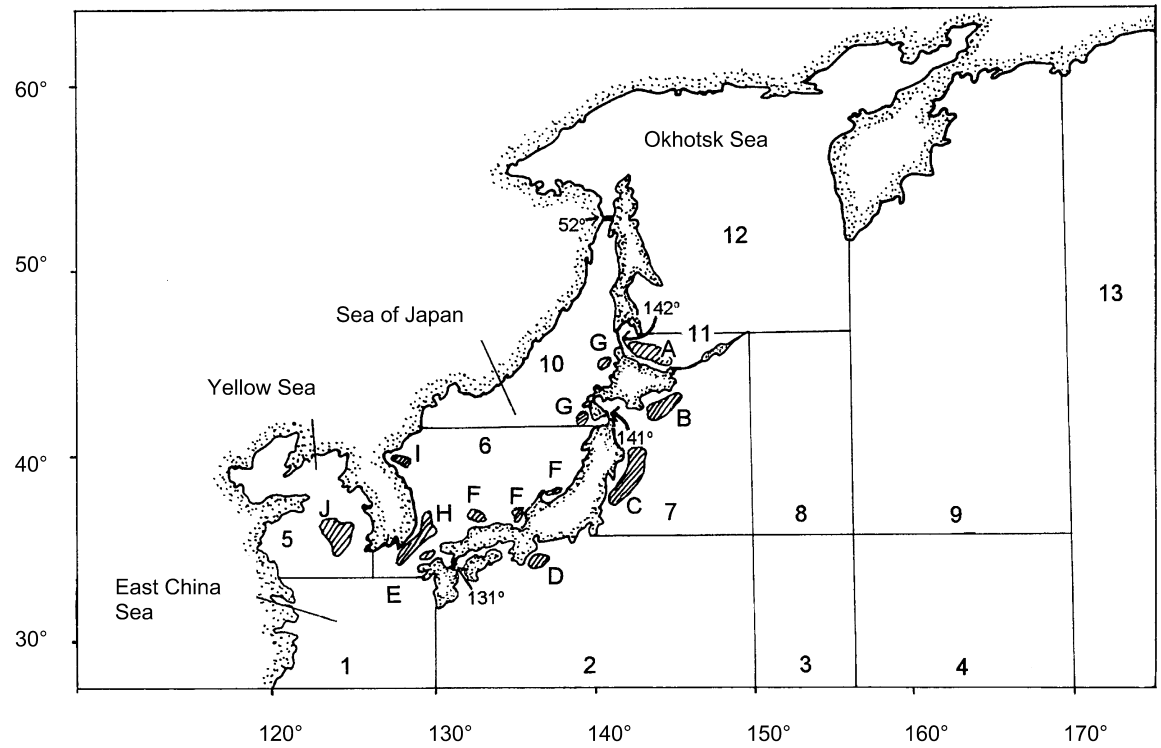


Fig. 2. Sub-areas of the western North Pacific as defined for minke whale management purposes (from IWC, 2000).

sub-area *i* is given by:

$$L(D^i \mid \underline{p}^i) = \frac{N^i!}{\prod_{j=1}^{k} x_j^i!} \prod_{j=1}^{k} (p_j^i)^{x_j^i} \quad (1)$$

where D^i is the dataset for sub-area *i*.

It is necessary to specify prior distributions for the parameters of model (1) to apply a Bayesian estimation approach and hence to compute Bayes Factors. The prior chosen for the parameters is the Dirichlet-β distribution, i.e.:

$$P(\underline{p}^i) = \frac{\Gamma\left(\sum_{j=1}^{k} \alpha_j\right)}{\prod_{j=1}^{k} \Gamma(\alpha_j)} \prod_{j=1}^{k} (p_j^i)^{\alpha_j - 1} \quad (2)$$

where $\underline{\alpha}$ are the parameters of the prior distribution.

The Dirichet-β distribution was chosen as the prior distribution because it is the conjugate prior to the multinomial distribution (Johnson and Kotz, 1970; Gelman *et al.*, 1995). It can be shown (Johnson and Kotz, 1970) that the marginal posterior distribution for sub-area *i* is given by:

$$\int_{\underline{p}^i} L(D^i \mid \underline{p}^i) P(\underline{p}^i \mid \underline{\alpha}) d\underline{p}^i = \frac{N^i!\, \Gamma(\sum_{j=1}^{k} \alpha_j) \prod_{j=1}^{k} \Gamma(x_j^i + \alpha_j)}{\prod_{j=1}^{k} x_j^i! \prod_{j=1}^{k} \Gamma(\alpha_j) \Gamma(\sum_{j=1}^{k} \alpha_j + N^i)} \quad (3)$$

The marginal posterior across all sub-areas is therefore the product over sub-areas of Equation (3), i.e.:

$$P^m(\mathbf{D} \mid \underline{\alpha}) = \prod_i \left(\frac{N^i!\, \Gamma(\sum_{j=1}^{k} \alpha_j) \prod_{j=1}^{k} \Gamma(x_j^i + \alpha_j)}{\prod_{j=1}^{k} x_j^i! \prod_{j=1}^{k} \Gamma(\alpha_j) \Gamma(\sum_{j=1}^{k} \alpha_j + N^i)} \right) \quad (4)$$

Fig. 3 explores the impact of different choices for the values for the parameters of the prior for the simple case in which there is only one sub-area and two haplotypes. For this example, the posterior can be summarised by the probability of getting one of the two haplotypes. The potential for the prior to 'bias' the posterior away from the probability implied by the data alone depends on (1) the sum of α over all haplotypes relative to the total sample size and (2) the relative difference between the ratio of the number of animals observed with each haplotype and the ratio of the αs. Furthermore, Equation (3) can be interpreted by noting that including the prior is equivalent to 'adding' a sample where the number of individuals with haplotype *j* is equal to $\alpha_j - 1$ to the actual data.

Now Equations (1)-(4) are based on the assumption that the proportion of animals in sub-area *i* with haplotype *j*, p_j^i, depends on sub-area, i.e. this is a multi-stock assumption. To develop the marginal posterior across all sub-areas for a single-stock model, the likelihood for the dataset for sub-area *i* and the prior are given by Equations (1) and (2) where the dependence of *p* on sub-area is dropped.

The marginal posterior across all sub-areas for the single stock model is given by:

$$P^s(\mathbf{D} \mid \underline{\alpha}) = \int_{\underline{p}} \prod_i L(D^i \mid \underline{p}) P(\underline{p} \mid \underline{\alpha})\, d\underline{p} \quad (5)$$

which can be shown to be:

$$\frac{\prod_i N^i!}{\prod_i \prod_{j=1}^{k} x_j^i!} \frac{\Gamma(\sum_{j=1}^{k} \alpha_j)}{\prod_{j=1}^{k} \Gamma(\alpha_j)} \frac{\prod_{j=1}^{k} \Gamma\left(\sum_i x_j^i + \alpha_j\right)}{\Gamma\left(\sum_{j=1}^{k} \alpha_j + \sum_i N^i\right)} \quad (6)$$

Dealing with the parameters of the prior

The specification of the values for the parameters of the Dirichlet-β prior (the αs) can be achieved using Empirical Bayes or Full Bayes approaches. The Empirical Bayes approach involves pre-specifying the values for the hyper-parameters (the αs) based on the actual data (i.e. the *x*s), while the Full Bayes approach involves placing a (hyper)prior[2] on the αs. One immediate difference between the Empirical and Full Bayes approaches is therefore that only the latter deals with the uncertainty associated with the αs. Pella and Masuda (2001) apply two methods ('maximum prior predictive distribution' and 'minimum squared-error risk') to determine Empirical Bayes estimates for the αs. In common with Pella and Masuda (2001), it was found here that the former method often leads to unrealistically large (i.e. very informative – see Fig. 3) values for the αs.

[2] A hyper-prior is the prior distribution for the parameters of the prior distribution.

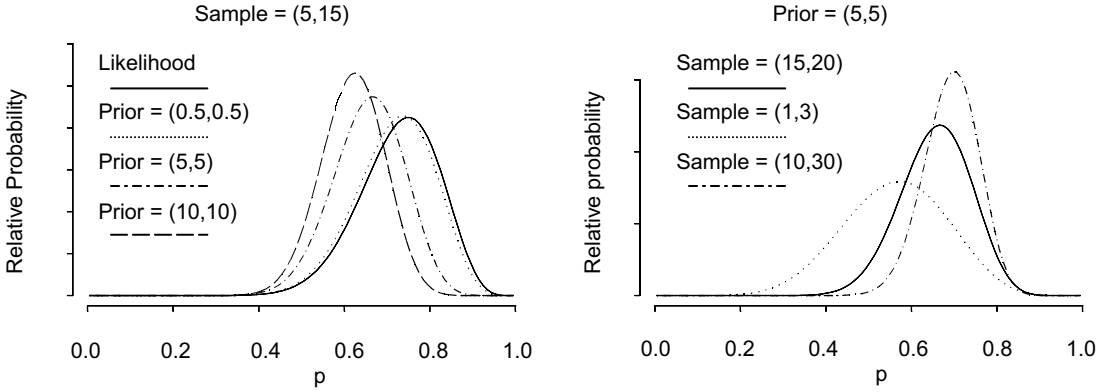


Fig. 3. Sensitivity of the posterior distribution for the probability of selecting an animal with a particular haplotype from a population where there are only two haplotypes. The likelihood is assumed to be binomial and the prior a beta distribution. The base-case specifications are that 15 animals out of a sample of 20 had the particular haplotype and the parameters of the beta distribution prior = (5,5).

Therefore, the Empirical Bayes results reported in this paper are based on the 'minimum squared-error risk' method (see Appendix 1 and Bishop *et al.* (1975) for details).

The hyper-prior for the Full Bayes approach needs to satisfy the constraint $\alpha_j > 0\ \forall j$. One widely-used distribution that automatically imposes this constraint is the log-normal distribution, i.e.:

$$P(\alpha_j) = \frac{1}{\sqrt{2\pi}\,\sigma_\alpha\,\alpha_j} e^{-\frac{1}{2\sigma_\alpha^2}(\ell n\alpha_j - \ell n\bar{\alpha})^2} d\alpha \qquad (7)$$

where:

$\bar{\alpha}$ is the median of the hyper-prior for α; and
σ_α is the standard deviation of the hyper-prior for $\ell n\ \alpha$.

The values for σ_α and σ_α are taken to be the mean and standard deviation respectively of the Empirical Bayes estimates for the $\ell n\ \alpha$. Therefore, the Full Bayes approach has been designed to be roughly comparable with the Empirical Bayes approach[3].

Computational aspects

Computing the Bayes Factor is straightforward if an Empirical Bayes approach is adopted because the Bayes Factor is simply the ratio of the marginal posteriors (Equations 4 and 6) given the values calculated for the αs. In contrast, the Full Bayes approach involves integrating over the hyper-prior for the αs; i.e., in this case, the Bayes Factor is defined as:

$$BF = \frac{\int_{\underline{\alpha}} P^s(\mathbf{D}\,|\,\underline{\alpha})\,P(\underline{\alpha})\,d\underline{\alpha}}{\int_{\underline{\alpha}} P^m(\mathbf{D}\,|\,\underline{\alpha})\,P(\underline{\alpha})\,d\underline{\alpha}} \qquad (8)$$

Evaluation of the numerator and the denominator of Equation (8) cannot be achieved analytically, and consequently a numerical integration approach needs to be applied. Three alternative approaches to computing the integrals were considered; two of these (Equations 9a and 9b) are based on samples from the posterior distribution for the αs (i.e. the distribution defined by $\frac{P^{s/m}(\mathbf{D}\,|\,\underline{\alpha})P(\underline{\alpha})}{\int P^{s/m}(\mathbf{D}\,|\,\underline{\alpha})P(\underline{\alpha})d\underline{\alpha}}$) while the third (Equation 9c) relies on samples from an approximation to that posterior distribution:

$$\hat{I} = \left\{ \frac{1}{w} \sum_{l=1}^{w} \frac{1}{P^{s/m}(\mathbf{D}\,|\,\underline{\alpha}_l)} \right\}^{-1} \qquad (9a)$$

$$\hat{I} = \left\{ \frac{1}{w} \sum_{l=1}^{w} \frac{f(\underline{\alpha}_l)}{P^{s/m}(\mathbf{D}\,|\,\underline{\alpha}_l)\,P(\underline{\alpha}_l)} \right\}^{-1} \qquad (9b)$$

$$\hat{I} = \frac{\sum_{l=1}^{q} P^{s/m}(\mathbf{D}\,|\,\underline{\alpha}_l)P(\underline{\alpha}_l)/g(\underline{\alpha}_l)}{\sum_{l=1}^{q} P(\underline{\alpha}_l)/g(\underline{\alpha}_l)} \qquad (9c)$$

where:

$\hat{I}$ is the integral needed to compute the Bayes factor;

w is the number of parameter vectors generated from the (numerical representation of the) posterior distribution;

q is the number of parameter vectors generated from $g()$, the approximation to the posterior distribution; and

$P^{s/m}(\mathbf{D}|\underline{\alpha}_l)$ is the value of the marginal posterior (see Equations 4 and 6 for the multiple and single stock marginal posteriors) for the lth parameter vector.

Equations (9a) and (9b), due to Newton and Raftery (1994) and Gelfand and Dey (1994) respectively, utilise a sample of w independent parameter vectors from the posterior distribution for the αs. Equation (9c), on the other hand, is based on a sample of q points from an approximation to this posterior distribution (the importance function). The function $f()$ in Equation (9b) can be any proper density function. For the purposes of this study, $f()$ and $g()$ have been taken to be the multivariate normal distribution with multivariate mean given by the vector $\underline{\alpha}_{\max}$ at which the marginal posterior attains its maximum and the variance-covariance matrix obtained by inverting the Hessian matrix about $\underline{\alpha}_{\max}$. The samples needed to apply Equations (9a)[4] and (9b) can be obtained using the Markov Chain Monte Carlo (MCMC) (Hastings, 1970; Gelman *et al.*, 1995) or Sample-Importance-Resample (SIR) (Rubin, 1987) algorithms, while those needed to apply Equation (9c) can be obtained using the SIR algorithm if the function $g()$ is taken to be the importance function.

A major problem associated with the application of Bayesian methods to complex problems is how to assess whether the algorithm used for numerical integration has converged to the posterior distribution (Gelman *et al.*, 1995). Assessing convergence can be divided into two parts in the context of this study: (a) whether the MCMC and SIR algorithms have been run for long enough that the resultant samples represent the posterior distributions; and (b) whether Equations (9a)-(9c) are numerically stable. The convergence of the SIR algorithm has been evaluated in this paper by the proportion of replicate parameter vectors in the sample from the posterior (typically no more then 0.5%), while the convergence of the MCMC algorithm has been evaluated using the magnitude of the correlation between 'adjacent' parameter sets (both visually and by means of correlation coefficients) and by using the statistics contained in the 'Bayesian Output Analysis' set of routines for assessing convergence of MCMC chains[5]. Based on these considerations, it was concluded that an adequate representation of the posterior could be obtained by conducting 5,570,000 cycles of the MCMC algorithm, ignoring the first 15% as a 'burn in' period, and then selecting every 2,000th parameter vector in the remaining chain. This resulted in a sample of 2,500 parameter vectors from the posterior distribution on which the integrals could be based. The SIR results are based on 5,000,000 draws from the importance function and 2,500 resamples from these draws.

The results in this paper are based on Equations 9b and 9c. Equation 9a was not used because it was found to be numerically unstable. The lack of stability of Equation (9a) is not surprising because a parameter vector with small likelihood can have a large impact on the value for $\hat{I}$ (Kass

[3] Basing the values for $\bar{\alpha}$ and σ_α on the Empirical Bayes estimates implies some use of the haplotype frequency data in developing the hyper-prior and, as such, the 'Fully Bayes' approach therefore has a slightly Empirical Bayes flavour.

[4] Equation (9a) follows from Equation (9c) taking as the sample from the importance function a sample from the posterior, i.e. $g(\underline{\alpha}_l) = P^{s/m}(\mathbf{D}|\underline{\alpha}_l)P(\underline{\alpha}_l)/\hat{I}$.

[5] http://www.pmeh.uiowa.edu/boa/.

and Raftery, 1995; Carlin and Louis, 2000). Table 1 illustrates this potential sensitivity for the simple case in which the prior is U[-5,5] and the likelihood is N(0; 1^2). For this case, it is possible to generate samples directly from the posterior distribution and to compute the integral over the product of the likelihood and prior analytically. While illustrative, the results in Table 1 indicate that improved performance arises from larger samples from the posterior distribution and that Equation 9a gives results which are much more variable than Equations 9b and 9c. It would appear that Equation 9a is only first order correct while Equations 9b and 9c are second order correct.

Example application

The *Implementation Simulation Trials* for the North Pacific minke whales (IWC, 2000; 2001a) include two hypotheses regarding stock structure in the western North Pacific: a two stock- ('J' and 'O') model and a three stock- ('J', 'O' and 'W') model[6]. There is support for at least two stocks in the Western North Pacific from analyses of allele frequency (Wada, 1984; 1991; Punt *et al.*, 1995; Butterworth *et al.*, 1996), conception date (Best and Kato, 1992), mtDNA (Goto and Pastene, 1997; 1998) and morphological (Kato *et al.*, 1992) information. The evidence for a 'W' stock is low *p*-values found in comparisons involving sub-area 9 (IWC, 2001b).

The mtDNA control region sequencing data used in this study were from minke whales taken during Korean and Japanese coastal small-type whaling operations (1982-1987; Goto and Pastene, unpublished data) and during the Japanese Whale Research Programme under Special Permit in the western North Pacific (JARPN) (1994-1999; Goto and Pastene, 2000). Some of the analyses conducted excluded samples from the western part of sub-area 9 (west of 162^0E) in 1995. The reason for doing this is that the results of previous hypothesis testing based on mtDNA data showed some heterogeneity in this particular group of animals (Goto *et al.*, 2000). Given these previous results, it was of interest to examine the sensitivity of the results from the Bayesian approach to including and excluding the data from the western part of sub-area 9 in 1995. The results of Goto *et al.* (2000) suggest the possibility of some temporal component to the distribution of stocks in the western North Pacific.

[6] More recent versions of these trials (IWC, 2001a) include the hypothesis that there is limited interchange between the 'O' and 'W' stocks but this hypothesis is not considered in this paper.

RESULTS AND DISCUSSION

Simulation evaluation

The objective of the method developed in this paper is that the resultant Bayes Factor should be very large if the one-stock hypothesis is correct, 0 if the two-stock hypothesis is correct and 1 if the data are unable to identify which stock structure hypothesis is correct. The ability of the method to achieve this objective can be evaluated by means of simulation (e.g. Martien and Taylor, 2000; Taylor *et al.*, 2000). Detailed simulations are, however, beyond the scope of the current paper. Nevertheless, some simulations have been conducted to evaluate the performance of the method given different sample sizes and true stock structure hypotheses.

Fig. 4 plots the logarithms of the ratio of the probability of the one-stock hypothesis to the sum of the probabilities of the one- and two-stock hypotheses (essentially the relative weight that should be assigned to the one-stock hypothesis) against the logarithms of the *p*-values from likelihood ratio tests comparing the one- and two-stock hypotheses[7]. The results in Fig. 4 are based on applying the Full Bayes version of the method to 20 datasets, each of which includes two areas. The haplotype frequency data for sub-areas 6 and 7 were used as the basis to generate the simulated data for the two areas; i.e. the two-stock hypothesis is correct for these simulations. As expected from previous studies, the one-stock hypothesis is rejected by both the Bayesian and frequentist methods even for low (25 per area) samples sizes. This result suggests that the method of this paper performs adequately when there are major differences in haplotype frequencies among areas.

Fig. 5 plots the relative weights that should be assigned to the one-stock hypothesis against the *p*-values from a likelihood ratio test comparing the one- and two-stock hypotheses for the case in which both datasets are generated from the haplotype frequency data for sub-area 7. This is a

[7] More powerful versions of the likelihood ratio test based on randomisation techniques exist, but the qualitative features of Figs 4–7 should be insensitive to using such tests.

Table 1

Sampled-based estimates of the integral of the product of a U[-5,5] prior and a N(0;1^2) likelihood. Results are shown for three alternative approximating equations, three alternative posterior sample sizes, and 10 individual samples from the posterior distribution for each sample size. The row 'SD' lists the standard deviation (multiplied by 100) of the 10 values for each combination of approximating equation and sample size.

Simulation number	Sample size/approximating equation								
	1,000,000			100,000			10,000		
	Eqn 9a	Eqn 9b	Eqn 9c	Eqn 9a	Eqn 9b	Eqn 9c	Eqn 9a	Eqn 9b	Eqn 9c
1	0.099	0.100	0.100	0.111	0.100	0.100	0.110	0.101	0.101
2	0.094	0.100	0.100	0.094	0.100	0.100	0.135	0.100	0.101
3	0.103	0.100	0.100	0.107	0.100	0.100	0.119	0.098	0.100
4	0.098	0.100	0.100	0.114	0.099	0.100	0.103	0.099	0.099
5	0.108	0.100	0.100	0.105	0.100	0.100	0.118	0.100	0.101
6	0.093	0.100	0.100	0.105	0.100	0.100	0.105	0.099	0.100
7	0.108	0.100	0.100	0.118	0.100	0.100	0.122	0.099	0.101
8	0.102	0.100	0.100	0.100	0.100	0.100	0.083	0.097	0.100
9	0.099	0.100	0.100	0.106	0.100	0.100	0.118	0.099	0.100
10	0.100	0.100	0.100	0.121	0.100	0.100	0.110	0.099	0.100
SD	0.523	0.021	0.015	0.820	0.033	0.021	1.378	0.105	0.072

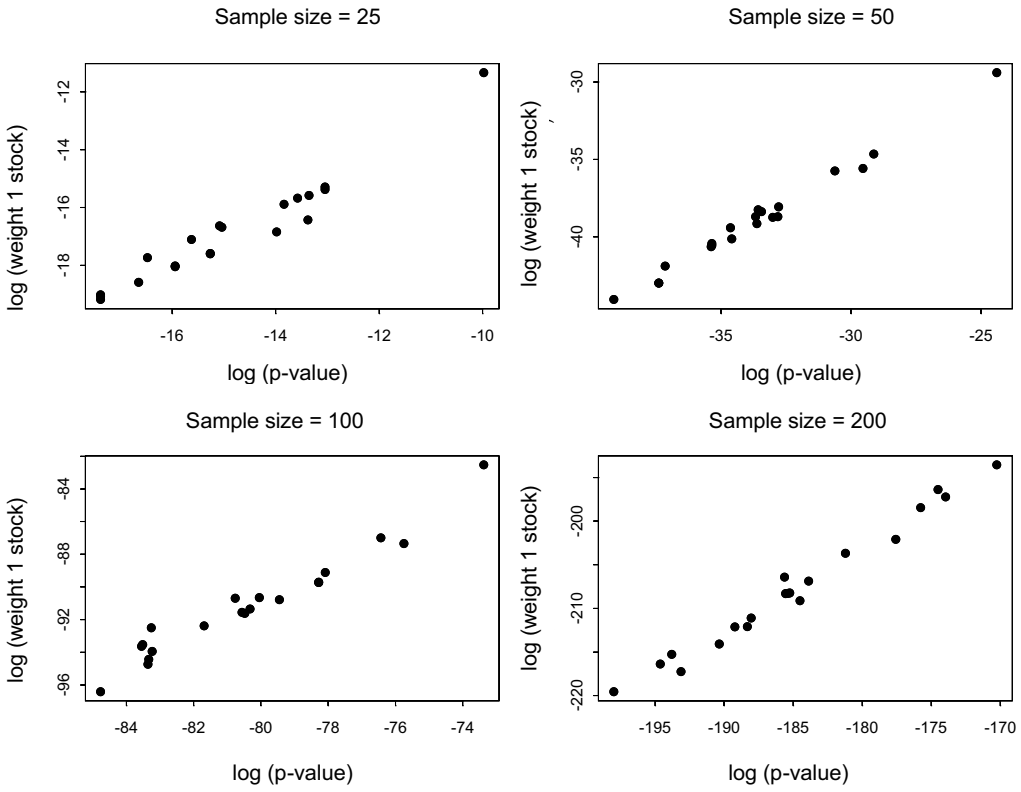


Fig. 4. Logarithms of the relative weights assigned to the one-stock hypothesis based on the Full Bayes method versus logarithms of likelihood ratio test *p*-values. The results in this Figure are based on 20 simulations in which the two-stock hypothesis is correct.

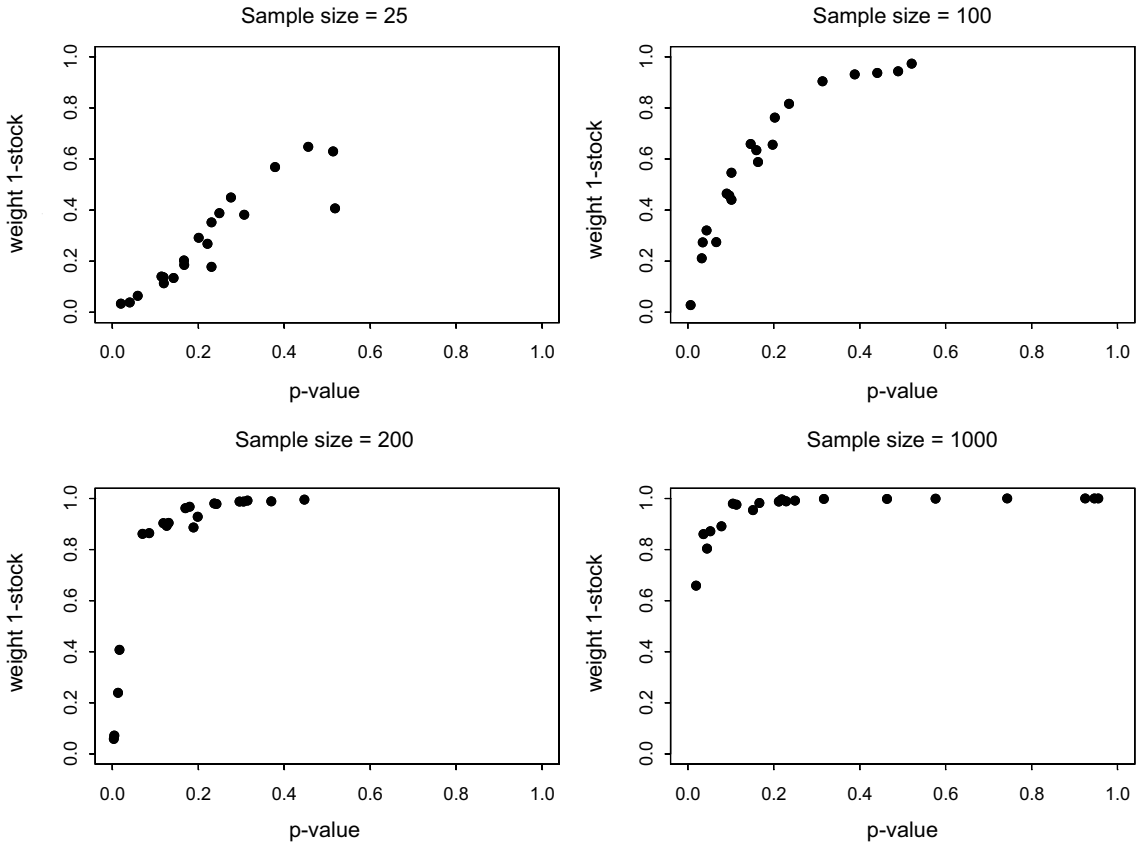


Fig. 5. Relative weights assigned to the one-stock hypothesis based on the Full Bayes method versus likelihood ratio test *p*-values. The results in this Figure are based on 20 simulations in which the one-stock hypothesis is correct.

case in which the one-stock hypothesis is correct. The results of the Full Bayes approach for the smallest sample size (25 per area) inappropriately indicate support for the two-stock hypothesis (17 of the 20 relative weights are smaller than 0.5) while three of the 20 likelihood ratio test *p*-values were smaller than the nominal level of 0.05. However, increasing the sample size from 25 to 100, to 200 and then to 1,000, results in much better performance, with increases in the

proportions of weights larger than 0.5 to 0.6, 0.8, and 1 respectively. In contrast, it is noteworthy that even for a sample size of 1000, three of the 20 simulations led to likelihood ratio test *p*-values less than the nominal level of 0.05. The reasons for the poor performance of the Full Bayes approach for a sample size of 25 are unclear but are probably related to the nature of the prior distribution, the effect of which is minimised given large sample sizes.

Although performance is adequate for large sample sizes, the results in Figs 4 and 5 suggest that the Full Bayes method may not provide reliable results for very small sample sizes if the one-stock hypothesis is correct. The simulations conducted to date are, however, relatively limited and additional work in this area is needed. Additional simulation work should be conducted in which the pseudo datasets are generated using either the types of operating models considered by Martien and Taylor (2000) or those based on coalescence simulations (e.g. Hudson, 1991). Both of these approaches to generating pseudo datasets allow these to be generated such that they are bounded by the limitations imposed by evolution and gene flow, although coalescence simulations are likely to be more efficient computationally.

Figs 6 and 7 show results analogous to those in Figs 4 and 5, except that results are shown for the Empirical Bayes approach (based on the 'minimum squared-error risk' method) and a variant of this approach in which the αs are set to 'uninformative' values (i.e. $\alpha = 1/k$ – Pella and Masuda (2001)). The performance of the Empirical Bayes approach (open symbols in Figs 6 and 7) is very poor for the case in which the one-stock hypothesis is correct (Fig. 6); even with very large sample sizes, the Empirical Bayes approach indicates a preference for the two-stock hypothesis. The performance of the 'uninformative' approach is markedly better than the Empirical and Full Bayes approaches when the one-stock hypothesis is correct (Figs 5 and 6). However, its performance is less than ideal for the case in which the two-stock hypothesis is correct (Fig. 7). Although the 'uninformative' approach places greatest weight on the two-stock hypothesis when it is correct, the weight assigned to the one-stock hypothesis is much greater than the *p*-value from the likelihood ratio test —this suggests that the 'uninformative' approach tends to 'favour' the one-stock hypothesis. The extent to which this is actually a concern is not entirely clear because even the 'uninformative' approach indicates that the two-stock hypothesis is far more likely than the one-stock hypothesis. A further problem with the 'uninformative' approach is that not all authors agree on the values to assign to the $\boldsymbol{\alpha}$s in order to obtain an 'uninformative' conjugate prior for a multinomial likelihood (Gelman *et al.*, 1995).

The only potential impediment to evaluating Bayes Factor approaches to addressing stock structure questions using simulation is their computational demands. This is not a problem with the Empirical Bayes or uninformative approaches as they do not involve any numerical integration. In contrast, this is certainly potentially a major problem for the Full Bayes approach. However, the software on which the calculations of this paper are based has been optimised so that roughly 10-20 Bayes Factors can be calculated in a twelve-hour period. This suggests that a full evaluation of even the Full Bayes approach should be feasible in short- to medium-term.

Illustrative application of the data for North Pacific minke whales

Table 2 lists Bayes Factors for a variety of comparisons among sub-areas for minke whales in the western North Pacific. Sub-area 7 comparisons are shown based on commercial samples, JARPN samples, and commercial and JARPN samples combined. The values for the Bayes Factor can be interpreted in terms of the support for (or against) the one-stock hypothesis (more positive values indicating

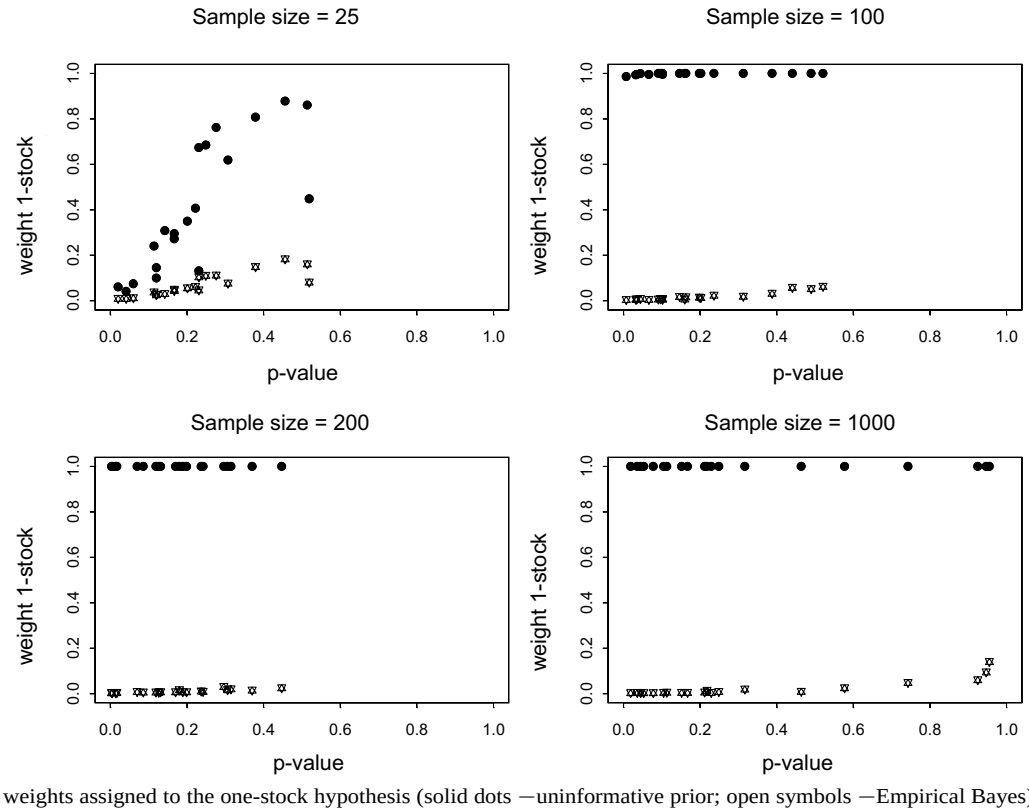


Fig. 6. Relative weights assigned to the one-stock hypothesis (solid dots —uninformative prior; open symbols —Empirical Bayes) versus likelihood ratio test *p*-values. The results in this Figure are based on 20 simulations in which the one-stock hypothesis is correct.

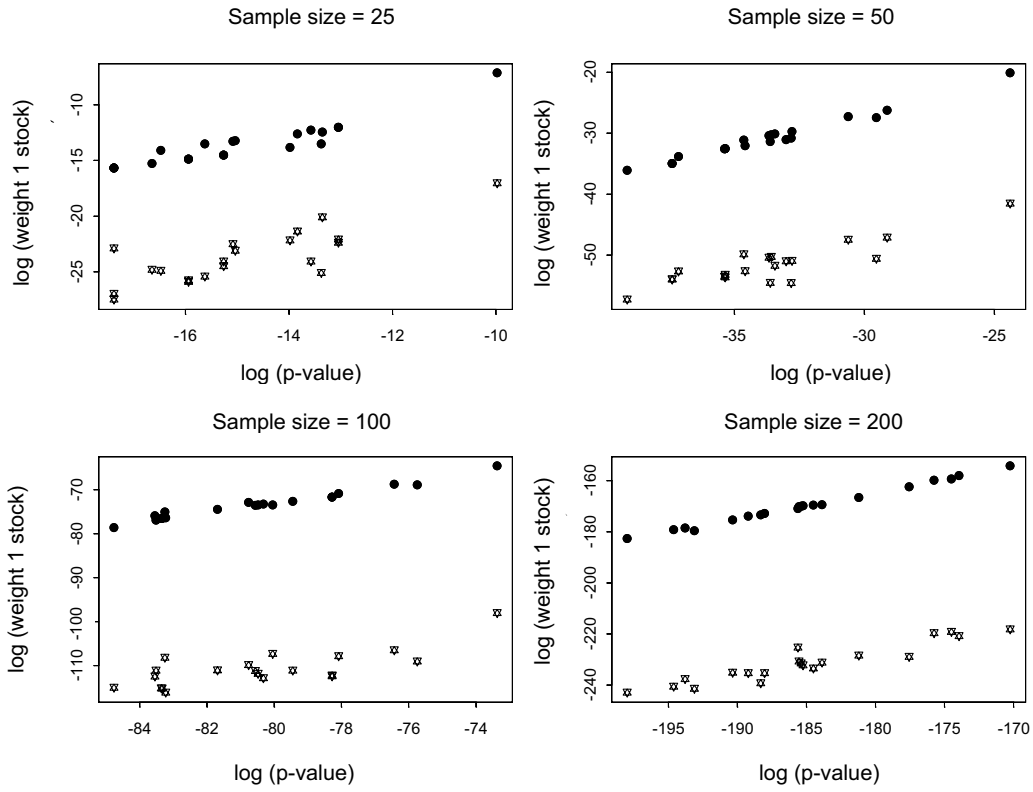


Fig. 7. Logarithms of the relative weights assigned to the one-stock hypothesis (solid dots —uninformative prior; open symbols —Empirical Bayes) versus logarithms of likelihood ratio test *p*-values. The results in this Figure are based on 20 simulations in which the two-stock hypothesis is correct.

greater support). Table 2 indicates the qualitative strength of evidence for the one-stock hypothesis using the scheme developed by Kass and Raftery (1995). Results are shown in Table 2 for the Full Bayes method (based on Equation 9c), and the 'uninformative' approach. Results are not shown for the Empirical Bayes approach given its poor performance in Fig. 6.

Account needs to be taken of the potential numerical uncertainty associated with the calculation of Bayes Factors using the Full Bayes approach when interpreting the results in Table 2. Fig. 8 shows the distribution for the Bayes Factor for a comparison of sub-areas 7 (JARPN and commercial samples combined) and 9 (less west 1995) that results from changing the random number sequence used when applying the MCMC and SIR algorithms. It is clear that the value of the Bayes factor can be sensitive to the random number sequence when Equation (9b) is used (it would be even more sensitive had the Bayes Factor been based on Equation 9a) while the results for Equation 9c are relatively insensitive to the random number sequence.

Table 2

Results for comparisons of some one- and two-stock hypotheses. The column 'Total haplotypes' lists the total number of haplotypes included in the analysis. The Bayes Factor is defined as the ratio of the probability of the one-stock hypothesis to that of the two-stock hypothesis. C = Commercial data; J = JARPN data; * = Positive; ** =Strong; and *** = Very strong.

Area 1	Area 2	Sample sizes		Total haplotypes	ℓn (Bayes Factor)	
					Uninformative.	Full Bayes
6	7 (C+J)	28	285	49	-31.89***	-46.02***
9 (less west 95)	7 (C+J) + 8	110	376	60	26.89***	5.18***
9 (less west 95)	7 (C+J)	110	285	52	20.28***	2.24*
9 (less west 95)	7 J	110	139	44	13.62***	1.21*
9 (less west 95)	7 C	110	146	45	10.87***	-0.35
8+9 (less west 95)	7(C+J)	201	285	60	41.50***	0.97
9	7 (C+J)+8	188	376	64	48.12***	1.28*
9	7 (C+J)	188	285	56	24.29***	-1.13*
9	7 J	188	139	50	27.01***	-0.63
9	7 C	188	146	50	20.44***	-1.92*
8+9	7 (C+J)	279	285	64	51.68***	-0.40
8	7 (C+J)	91	285	56	13.78***	-1.56*
7C	7J	146	139	46	13.08***	0.85

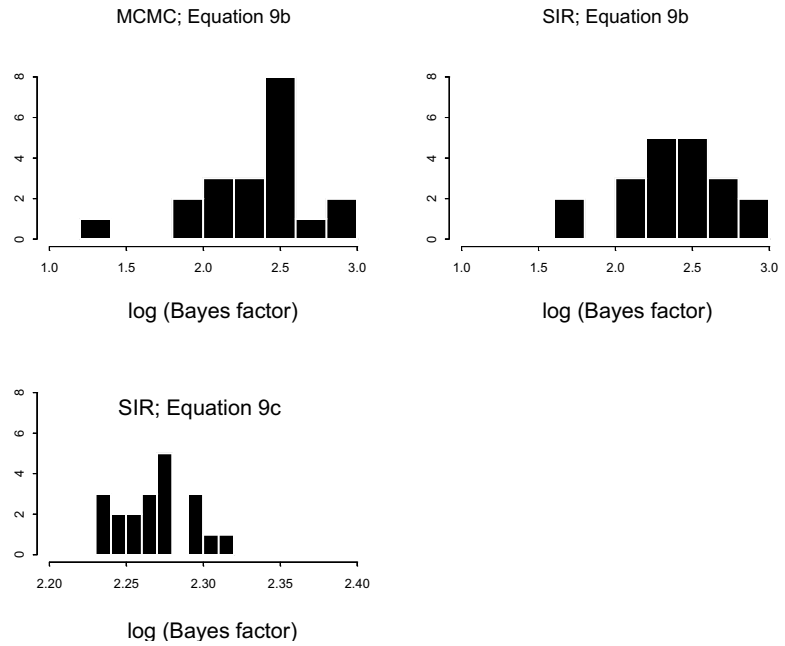


Fig. 8. Sensitivity of the logarithm of the Bayes Factor (probability one-stock hypothesis divided by probability two-stock hypothesis) to the choice of the random number sequence used when conducting the Monte Carlo integrals and the equation used to approximate the probability of the data. The results in this Figure relate to a comparison of sub-areas 7 (JARPN and commercial samples combined) and 9 (less west 1995).

The results, except for the comparison of sub-areas 6 and 7, are very sensitive to how the αs are specified. For example, there is very strong evidence for the one-stock hypothesis if an 'uninformative' prior is assumed for the αs. This may, however, be a consequence of the 'uninformative' approach tending to prefer the one-stock hypothesis. Nevertheless, it should be borne in mind that it is hard to extrapolate from the results in Fig. 7 (where the two-stock hypothesis is clearly correct) to the situation to the east of Japan.

The results for the Full Bayes approach suggest that the data generally provide little information regarding comparison among sub-areas 7, 8 and 9. The only marked exception to this is the comparison between sub-area 9 (less west 95) and the data for sub-areas 7 and 8 pooled. This result is consistent with those from a Bayesian analysis of the allele frequency data (Punt *et al.*, 2000) as well as with those from previous analyses based on hypothesis testing using allozymes (Wada, 1984) and mtDNA RFLP data (Goto and Pastene, 1997). For all combinations of factors, the probability that sub-area 9 contains a separate stock from sub-area 7 decreases if the data for the west of sub-area 9 in 1995 are omitted. This result is consistent with previous analyses based on hypothesis testing (Goto *et al.*, 2000).

DISCUSSION

The use of Bayesian methods to analyse genetics data, while preferable theoretically, is a relatively recent development. A prime reason for this is that the computational requirements of the calculations can be prohibitive. However, the opportunities for using these methods should increase with the advent of faster personal computers. Some of the assumptions made in this paper (for example that the prior for the proportion of animals in a sub-area with a given haplotype is of the Dirichlet-β form) were made largely for computational convenience (so that the marginal posterior could be evaluated analytically) rather than for good theoretical reasons. Although examining different choices for this prior is beyond the scope of the current study, development of more powerful computers should enable this to be carried out in the future.

The use of the log-normal hyper-prior for the αs is relatively arbitrary. It would seem prudent to examine the sensitivity of any results to be used for management purposes to other probability distribution functions that have similar properties to the log-normal (e.g. the gamma distribution). In addition, basing the mean and coefficient of variation of the hyper-prior on the values for the αs used in the Empirical Bayes calculations is also relatively arbitrary. However, no more objective way to define these parameters is immediately obvious. It should also be noted that the illustrative simulations only considered as the case when the two-stock hypothesis is correct, an example where there are clear differences in haplotype frequencies. Future simulations should consider scenarios in which the differences are less clear.

An advantage of Bayesian over frequentist approaches is that the former can be used to assign probabilities to alternative hypotheses. This cannot be achieved using frequentist techniques *inter alia* because the 'effect size' is unknown. This problem is not removed through the Bayesian approach. In fact, the 'effect size' is implicit in the priors. Fig. 9 illustrates the 'effect size' in terms of the prior distribution implied for the *Fst* statistic under the one- and two-stock models (solid and dotted lines respectively). Results are shown in Fig. 9 for two different choices for σ_α (0.5 and 1.5) and two choices for the sample sizes from the two sub-areas. As expected, there is considerable overlap between these distributions, particularly for the lower sample sizes and the high value for σ_α. High values for σ_α imply a more skewed haplotype frequency distribution (a few very common haplotypes and many rare haplotypes) due to the prior assigning higher probability to occasional large values for α. The distributions in Fig. 9 raise the intriguing question of whether a Bayesian analysis could be based on the implied prior distribution for a quantity such as *Fst* rather than having to be based on the Full Bayesian analysis.

At the present stage of development, the results from Full Bayes method should only be used for management purposes

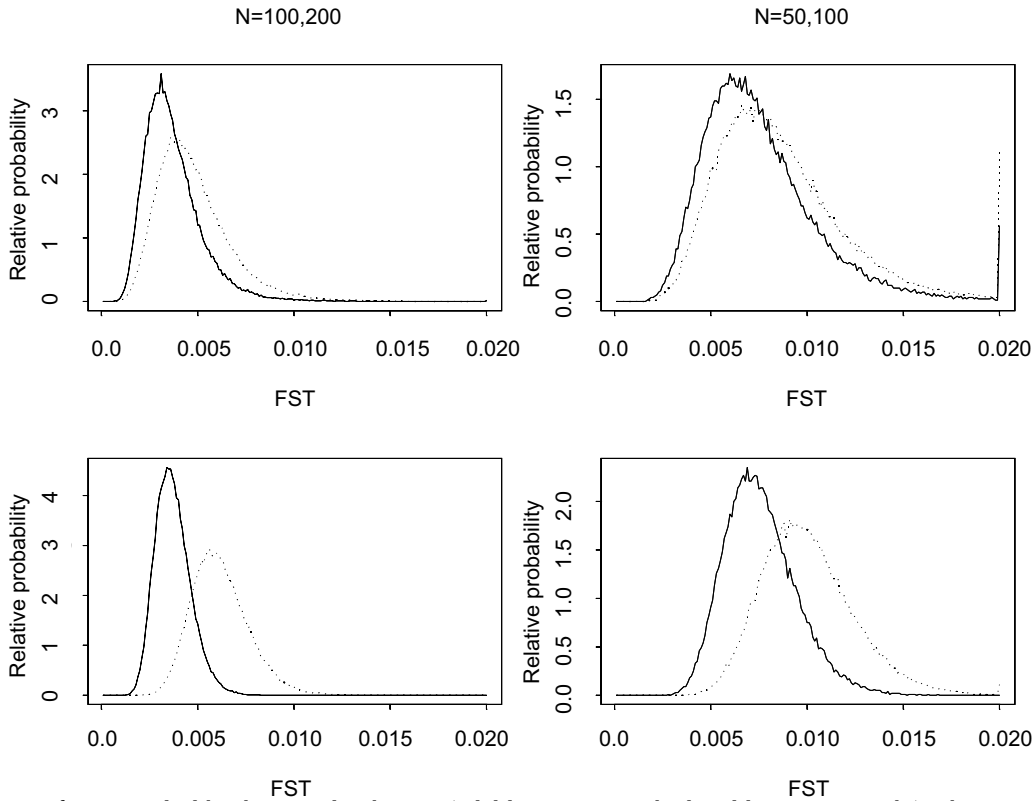


Fig. 9. Distributions for *Fst* implied by the prior distributions (solid lines —one stock; dotted lines —two stocks). The upper and lower panels correspond to σ_α = 1.5 and σ_α = 0.5 respectively. The left and right panels correspond to different sample sizes for the two areas: (100, 200) and (50, 100) respectively.

with caution. Some testing of the method has been completed and this suggests that, when applied to data for two stocks that differ substantially in haplotype frequency, or when sample sizes are large and the one-stock hypothesis is correct, performance is adequate. However, in common with most other methods for analysing genetics data, performance of this method has yet to be fully evaluated, particularly for cases in which there are two stocks but their haplotype frequencies differ only slightly.

Although the use of Bayesian methods for resolving stock structure questions is still in its infancy, we believe that these methods show considerable promise. For example, Pella and Masuda (2001) developed an approach based on similar assumptions regarding the likelihood function and the prior for the proportion of animals with a particular haplotype to estimate probability distributions for stock mixture rates. Further development of the technique outlined in this paper should provide a firmer basis for the development of *Implementation Simulation Trials*. Conditioning of *Implementation Simulation Trials*[8] is, in some cases, already based on a Bayesian assessment (e.g. Punt and Smith, 1999; IWC, 2002). In fact, there is no reason (barring computational constraints) therefore that genetics data could not be included in the conditioning process so that the probability of alternative stock structure hypotheses is one outcome of this process.

ACKNOWLEDGEMENTS

We would like to thank the Institute of Cetacean Research for allowing us to use the mtDNA data for North Pacific minke whales for the illustrative calculations of this paper. Doug Butterworth, Tony Smith, Bob Ward, the members of the 2001 International Whaling Commission Sub-Committee on Stock Definition and two anonymous reviewers are thanked for their comments on an earlier version of the paper.

[8] Conditioning involves selecting the values for the parameters of the operating model that represents the 'true situation' for the trials so that this model adequately mimics the observed data.

REFERENCES

Best, P.B. and Kato, H. 1992. Report of the sub-committee on North Pacific minke whales, Appendix 5. Possible evidence from foetal length distributions of the possible mixing of different components of the Yellow Sea-East China Sea-Sea of Japan-Okhotsk Sea minke whale population(s). *Rep. int. Whal. Commn* 42:166.

Bishop, Y.M.M., Fienberg, S.E. and Holland, P.W. 1975. *Discrete Multivariate Analysis: Theory and Practice*. MIT Press, Cambridge, MA. 557pp.

Bossart, J.L. and Pashley Powell, D. 1998. Genetic estimates of population structure and gene flow: limitations, lessons and new directions. *Trends Ecol. Evol.* 13:202-6.

Butterworth, D.S., Geromont, H.F. and Wada, S. 1996. Further analyses of allele frequency data to provide estimates of the extent of mixing between the various North Pacific minke whale stocks, and their implications for the status of the 'O' stock. *Rep. int. Whal. Commn* 46:443-51.

Carlin, B.P. and Louis, T.A. 2000. *Bayes and Empirical Bayes Methods for Data Analysis*. 2nd Edn. Chapman and Hall/CRC, Boca Raton. xvii+419pp.

Dizon, A.E., Perrin, W.F., Amos, W., Baker, C.S., Chivers, S.J., Costa, A.S., Curry, B.E., Gaggiotti, O., Hoelzel, A.R., Hofman, R., LeDuc, R.G., Loughlin, T.R., Lux, C.R., O'Corry-Crowe, G.M., Rosel, P.E., Rosenberg, A., Scribner, K.T. and Taylor, B.L. 1997. Report of the Workshop on the Analysis of Genetic Data to Address Problems of Stock Identity as Related to Management of Marine Mammals. pp. 3-48. *In:* A.E. Dizon, S.J. Chivers and W.F. Perrin (eds.) *Molecular Genetics of Marine Mammals*. The Society for Marine Mammalogy, Lawrence, KS.

Excoffier, L., Smouse, P.E. and Quattro, J.M. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131:479-91.

Fujise, Y. 2000. Outline of the research activities of the Japanese Whale Research Program under Special Permit in the North Pacific

(JARPN) from 1994 to 1999. Paper SC/F2K/J8 presented at the JARPN Review Meeting, Tokyo, Japan, 7-10 February 2000 (unpublished) 30pp. [Available from the Office of this Journal].
Gardner, M.G. and Ward, R.D. 1998. Population structure of the Australian gummy shark (*Mustelus antarcticus* Günther) inferred from allozymes, mitochondrial DNA and vertebrae counts. *Mar. Freshw. Res.* 49:733-45.
Gelfand, A.E. and Dey, D.K. 1994. Bayesian model choice: asymptotics and exact calculations. *J. Roy. Statist. Soc.* 56:501-14.
Gelman, A., Carlin, B.P., Stern, H.S. and Rubin, D.B. 1995. *Bayesian Data Analysis*. Chapman & Hall, London. [xix]+526pp.
Goto, M. and Pastene, L.A. 1997. Population structure of the western North Pacific minke whale based on an RFLP analysis of the mitochondrial DNA region. *Rep. int. Whal. Commn* 47:531-7.
Goto, M. and Pastene, L.A. 1998. Population structure in the North Pacific minke whale as revealed by RFLP and sequencing analyses of the mtDNA control region. Paper SC/50/RMP7 presented to the IWC Scientific Committee, April 1998 (unpublished). 15pp. [Paper available from the Office of this Journal].
Goto, M. and Pastene, L.A. 2000. Population structure in the western North Pacific minke whale based on RFLP and sequencing analyses of mtDNA control region, using data from 1994-1999 JARPN surveys. Paper SC/F2K/J11 presented at the JARPN Review Meeting, Tokyo, Japan, 7-10 February 2000 (unpublished). 16pp. [Paper available from the Office of this Journal].
Goto, M., Abe, H. and Pastene, L.A. 2000. Additional analyses of mtDNA control region sequences in the western North Pacific minke whales using JARPN samples. Paper SC/F2K/J32 presented at the JARPN Review Meeting, Tokyo, Japan, 7-10 February 2000 (unpublished) 5pp. [Paper available from the Office of this Journal].
Hall, M.A. and Donovan, G.P. 2001. Environmentalists, fishermen, cetaceans and fish: is there a balance and can science help to find it? pp. 491-521. *In:* P.G.H. Evans and J.A. Raga (eds.) *Marine Mammals: Biology and Conservation*. Kluwer Academic/Plenum Publishers, New York.
Hastings, W.K. 1970. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* 57:97-109.
Hudson, R.R. 1991. Gene genealogies and the coalescent process. *Oxf. Surv. Evol. Biol.* 9:1-44.
Hudson, R.R., Boos, D.D. and Kaplan, N.L. 1992. A statistical test for detecting geographic subdivision. *Mol. Biol. Evol.* 9:138-51.
International Whaling Commission. 1991. Report of the Workshop on the Genetic Analysis of Cetacean Populations, La Jolla, 27-29 September 1989. *Rep. int. Whal. Commn* (special issue) 13:3-21.
International Whaling Commission. 1992. Report of the Third Comprehensive Assessment Workshop on Management Procedures, Oslo, 1-8 February 1990. *Rep. int. Whal. Commn* 42:271-304.
International Whaling Commission. 1994. Report of the Scientific Committee. Annex H. The Revised Management Procedure (RMP) for Baleen Whales. *Rep. int. Whal. Commn* 44:145-52.
International Whaling Commission. 2000. Report of the Scientific Committee. Annex D. Report of the Sub-Committee on the Revised Management Procedure. *J. Cetacean Res. Manage. (Suppl.)* 2:79-92.
International Whaling Commission. 2001a. Report of the Scientific Committee. Annex D. Report of the Sub-Committee on the Revised Management Procedure. *J. Cetacean Res. Manage. (Suppl.)* 3:90-125.
International Whaling Commission. 2001b. Report of the Workshop to Review the Japanese Whale Research Programme under Special Permit for North Pacific Minke Whales (JARPN), Tokyo, 7-10 February 2000. *J. Cetacean Res. Manage. (Suppl.)* 3:377-477.
International Whaling Commission. 2002. Report of the Third Workshop on the Development of an Aboriginal Subsistence Whaling Management Procedure (AWMP). *J. Cetacean Res. Manage. (Suppl.)* 4:415-53.
Jeffreys, H. 1961. *Theory of Probability*. 3rd Edn. Oxford University Press, London. i-vii+411pp.
Johnson, N. and Kotz, S. 1970. *Distributions in Statistics: Discrete Distributions*. John Wiley & Sons, New York. xiv +328pp.
Kass, R.E. and Raftery, A.E. 1995. Bayes factors. *J. Am. Stat. Assoc.* 90:773-95.
Kato, H., Kishiro, T., Fujise, Y. and Wada, S. 1992. Morphology of minke whales in Okhotsk Sea, Sea of Japan and off the East coast of Japan, with respect to stock identification. *Rep. int. Whal. Commn* 42:437-42.
Kitada, S., Hiyashi, T. and Kishiro, H. 2000. Empirical Bayes procedures for estimating genetic distance between populations and effective population size. *Genetics* 156:2063-79.
Lulhart, G. and England, P.R. 1999. Statistical analysis of microsatellite DNA data. *Trends Ecol. Evol.* 14:253-56.
Martien, K.K. and Taylor, B.L. 2000. The limitations of hypothesis testing as a means of demographically delineating independent units. Paper SC/F2K/J3 presented at the JARPN Review Meeting, Tokyo, Japan, 7-10 February 2000 (unpublished). 32pp. [Paper available from the Office of this Journal].
Newton, M.A. and Raftery, A.E. 1994. Approximate Bayesian inference by the weighted likelihood bootstrap (with discussion). *J. Roy. Statist. Soc.* 56:1-48.
Pella, J. and Masuda, M. 2001. Bayesian methods for analysis of stock mixtures from genetic characters. *Fish. Bull.* 99:151-67.
Punt, A.E. and Smith, A.D.M. 1999. Harvest strategy evaluation for the eastern stock of gemfish (*Rexea solandri*). *ICES J. Mar. Sci.* 56(6):860-75.
Punt, A.E., Butterworth, D.S., Wada, S. and Kishino, H. 1995. On the use of allele frequency data to provide estimates of the extent of mixing between the various North Pacific minke whale stocks. *Rep. int. Whal. Commn* 45:293-300.
Punt, A.E., Butterworth, D.S. and Wada, S. 2000. On the use of allele frequency data within a Bayesian framework to evaluate the relative probabilities of alternative stock structure hypotheses for the North Pacific minke whales. *J. Cetacean Res. Manage.* 2(2):151-8.
Rubin, D.B. 1987. Comment on 'The calculation of posterior distributions by data augmentation'. *J. Am. Stat. Assoc.* 82:543-6.
Shoemaker, J.S., Painter, I.S. and Weir, B.S. 1999. Bayesian statistics in genetics: A guide for the uninitiated. *Trends in Genetics* 15:354-8.
Taylor, B.L., Chivers, S.J., Sexton, S. and Dizon, A.E. 2000. Using simulation models that incorporate uncertainty to estimate dispersal rates from mitochondrial DNA data. *Conserv. Biol.* In Press
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. 1991. Genetic structure of Okhotsk Sea-West Pacific stock of minke whales. Paper SC/43/Mi32 presented to the IWC Scientific Committee, May 1991 (unpublished). 17pp. [Paper available from the Office of this Journal].

Appendix 1

The minimum squared-error risk approach to selecting values for the αs

The values for the αs are defined according to the formula:

$$\alpha_j = \alpha^T \bar{x}_j \tag{1.1}$$

where:

α^T is a value that minimises the expected squared-error between the posterior means for the relative frequencies of each haplotype and the observed relative frequency of each haplotype; and

$\bar{x}_j$ is the arithmetic average of the relative frequency of haplotype *j* across samples:

$$\bar{x}_j = \frac{1}{n}\sum_{i=1}^{n}\frac{x_j^i}{N^i} \tag{1.2}$$

Now, it can be shown (Bishop *et al.*, 1975) that the value of α^T satisfies the equation:

$$\alpha^T \sum_{i=1}^{n}\left\{\frac{(N^i)^2}{(N^i+\alpha^T)^3}\sum_{j=1}^{k}(x_j^i-\bar{x}_j)^2\right\} = \sum_{i=1}^{n}\left\{\frac{(N^i)^2}{(N^i+\alpha^T)^3}[1-\sum_{j=1}^{k}(x_j^i)^2]\right\} \tag{1.3}$$

Solving Equation (1.3) for α^T and applying Equation (1.1) provides the 'minimum squared-error risk' values for the αs.

Population identity of humpback whales (*Megaptera novaeangliae*) in the waters of the US mid-Atlantic states

SUSAN G. BARCO[*+], WILLIAM A. MCLELLAN[#], JUDITH M. ALLEN[¥], REGINA A. ASMUTIS-SILVIA[++], RICHARD MALLON-DAY[**], ERIN M. MEAGHER[#], D. ANN PABST[#], JOOKE ROBBINS[##], ROSEMARY E. SETON[¥], W. MARK SWINGLE[*], MASON T. WEINRICH[±] AND PHILLIP J. CLAPHAM[=]

Contact e-mail: ocrab@erols.com

ABSTRACT

In recent years, humpback whales (*Megaptera novaeangliae*) have been observed in the waters of the US mid-Atlantic states (USMA; New Jersey to North Carolina), notably in winter. The level of the mortality in this area (52 recorded deaths from 1990-2000), makes it important to understand the nature and population identity of this aggregation. Of the approximately 100 humpback whales documented in this study, photographs of 41 (live or dead) were of sufficient quality to be compared to catalogues from the Gulf of Maine (GOM, the closest feeding ground) and elsewhere in the North Atlantic. Of 22 live whales, 10 (45.5%) matched to the GOM, 5 (22.7%) to Newfoundland and 1 (4.5%) to the Gulf of St Lawrence (GSL). Of 19 dead whales, 6 (31.6%) were known GOM whales. Although the population composition of the USMA is dominated by GOM whales, lack of recent photographic effort in Newfoundland makes it likely that the observed match rates under represent the true presence of Canadian whales in the region. Length data from 48 stranded whales (18 females, 22 males and 8 of unknown sex) suggest that 39 (81.2%) were first-year animals, 7 (14.6%) were immature and 2 (4.2%) were adults. However, sighting histories of five of the dead whales indicate that some were small for their age and histories of live whales further indicate that the population contains a greater percentage of mature animals than is suggested by the stranded sample. The authors suggest that the study area primarily represents a supplemental winter feeding ground that is used by humpbacks for more than one purpose. From a management perspective, although the only successful matches of mortalities to date have been to the GOM, the observed mixing of live whales from different summer stocks might suggest that the high numbers of mortalities occurring there may not be impacting this single stock alone. Although further data are required before conclusions can be drawn, the mortality rate may be significant for the GOM population and this warrants further investigation.

KEYWORDS: HUMPBACK WHALE; NORTH ATLANTIC; POPULATION IDENTITY; MIGRATION, MORTALITY; STRANDINGS

INTRODUCTION

In summer, humpback whales (*Megaptera novaeangliae*) in the North Atlantic are distributed from the eastern coast of the USA north to the waters of the Arctic (Clapham and Mead, 1999; Smith *et al.*, 1999; IWC, 2002). Individual whales show strong fidelity to specific feeding grounds within this range, including the Gulf of Maine, Newfoundland/Labrador, the Gulf of St Lawrence, Greenland, Iceland and Norway (Katona and Beard, 1990; Smith *et al.*, 1999). This fidelity is maternally directed (Clapham and Mayo, 1987) and in some areas persists for long enough to be reflected in the genetic structure of the population (Larsen *et al.*, 1996). Despite this strong segregation, whales from all of the known feeding grounds migrate to a common winter breeding range in the West Indies, where they mate and calve (Palsbøll *et al.*, 1997). It is generally believed that the majority of whales engage in this seasonal migration; however humpbacks have also been observed at high latitudes during winter (Charif *et al.*, 2001; J. Robbins, unpublished data).

In recent years, the occurrence of humpback whales has been documented from the coastal waters of the US mid-Atlantic states (hereafter USMA) from New Jersey to North Carolina (e.g. Swingle *et al.*, 1993; Wiley *et al.*, 1995). Most records are for the January to April period, although occasional sightings are made in the summer. Recorded mortalities are relatively high in the USMA compared to other coastal waters. Between 1990 and 2000 there were 52 reported deaths (Wiley *et al.*, 1995; Waring *et al.*, 2000). Of these, the cause of death could not be determined for 39 (75%). For the remaining 13, 11 were identified as being due to entanglements in fishing gear or collisions with vessels, whereas two showed no signs of human-induced mortality.

Knowledge of the origin of the USMA animals is important since managers must assign anthropogenic mortalities to the correct management stock in order to assess status. The closest feeding aggregation to the USMA is the well-studied Gulf of Maine aggregation that is considered a separate management unit by the US National Marine Fisheries Service (NMFS; Waring *et al.*, 2000). The implications for management will differ if all mortalities in the USMA are from this aggregation or from two (or more) aggregations.

In order to investigate the origin of the USMA animals, identifying photographs of living and dead USMA humpback whales were compared to catalogues from the Gulf of Maine, Newfoundland and other areas of the North

[*] *Virginia Marine Science Museum, 717 General Booth Boulevard, Virginia Beach, VA 23451, USA.*
[+] *Program in Ecological Sciences, Old Dominion University, Hampton Boulevard, Norfolk, VA 23529, USA.*
[#] *University of North Carolina at Wilmington, Biological Sciences and Center for Marine Science, 601 South College Road, Wilmington, NC 28403, USA.*
[¥] *College of the Atlantic, 105 Eden Street, Bar Harbor, ME 04609, USA.*
[++] *International Wildlife Coalition, 70 East Falmouth Highway, East Falmouth, MA 02536, USA.*
[**] *Cape May County Dolphin Survey, 40 Orchard Lane, Berwyn, PA 19312, USA.*
[##] *Center for Coastal Studies, Box 1036, Provincetown, MA 02657, USA.*
[±] *Whale Center of New England, Box 159, Gloucester, MA 01930, USA.*
[=] *Northeast Fisheries Science Center, 166 Water Street, Woods Hole, MA 02543, USA.*

Atlantic. In addition, patterns of occurrence were combined with available life history data of matched individuals to gain insight into the question of habitat use by humpback whales in the USMA.

MATERIALS AND METHODS

The USMA study area is shown in Fig. 1.

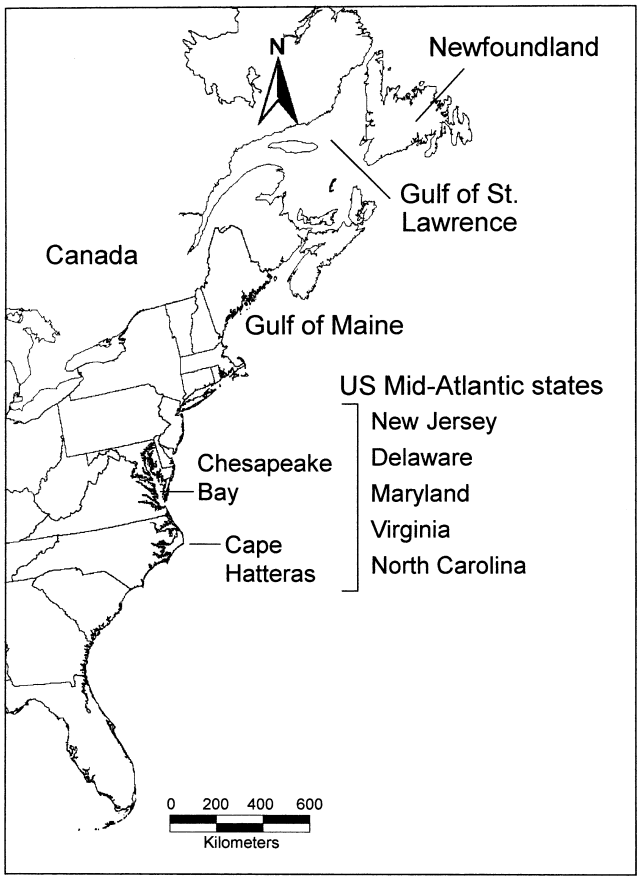


Fig. 1. Study area, in relation to other major regions mentioned in this report.

Observations of live humpback whales derive from a combination of sporadic vessel surveys (particularly in 1993 and 1994) directed at this or other species in the study area and whalewatching excursions based in Virginia Beach, VA and Cape May, NJ. Data collected included date, time, location and behaviour; where possible, photographs were taken for the purpose of individual identification. In the case of dead animals, information was collected by various researchers who responded to the stranding events. Where possible, data on the length and sex of the animal were obtained.

All available photographs of humpback whales observed either alive or dead in this area were included in this study. Most of the images were collected in the Mid-Atlantic Humpback Whale Catalogue (MAMNC) curated by the Virginia Marine Science Museum in Virginia Beach, VA (Swingle and Barco, 2000). Individual humpback whales were identified and catalogued using variations in the ventral fluke pattern and/or in the shape, size and scarring of the dorsal fin (Katona and Whitehead, 1981). These samples include, but are not limited to, those described in Wiley *et al.* (1995) and Swingle *et al.* (1993).

Each fluke photograph was compared to two large ocean-wide catalogues: the North Atlantic Humpback Whale Catalogue (NAHWC) and the collection from the Years of the North Atlantic Humpback (YONAH) project. The NAHWC includes photographs of 5,341 individuals, with observations dating from 1968 to the present. The YONAH collection includes 2,998 individuals, with the great majority photographed during a two-year ocean-basin-wide collaborative study in 1992 and 1993 (Smith *et al.*, 1999). Both catalogues include photographs from all studied areas of the North Atlantic, including the Gulf of Maine, the Gulf of St Lawrence, Newfoundland, Labrador, West Greenland, Iceland, Norway and the West Indies. Both catalogues are curated by the College of the Atlantic in Bar Harbor, Maine.

In addition to the oceanic catalogues described above, photographs in this study were compared to two regional Gulf of Maine catalogues. While these catalogues overlapped with the oceanic collections to some degree, they contained more recent coverage and more detailed information on the individual animals in that region. Regional catalogues also maintained photographic coverage of the dorsal fin that was an additional source of potential matches to USMA photographs. Gulf of Maine researchers collaborate to identify new additions to the population, but maintain separate photographic catalogues and archives that differ in their underlying levels of effort and geographic coverage. In this study, comparisons were made to a catalogue curated by the Center for Coastal Studies (CCS) in Provincetown, Massachusetts and to one curated by the Whale Center of New England (WCNE) in Gloucester, Massachusetts. The CCS catalogue includes photos of 1,273 individuals sighted from 1975 to the present, including directed effort throughout the Gulf of Maine. The WCNE collection includes 1,419 individuals, with the earliest observations from 1974.

Animals deemed less likely to be successfully matched to high latitude catalogues due to low image quality or animal distinctiveness were excluded from the matching effort with the NAHWC and GOM catalogues. In the case of stranding documentation, the condition of the carcass was also considered. However, all available photographic documentation, regardless of feature or quality, was used to establish within- and between-year resightings of individuals in the USMA.

Exact or minimum ages of USMA animals were determined based on previous sightings in other regions. An exact age was known for animals first catalogued as calves. Animals without a known year of birth were assigned a minimum age by assuming that the whale was at least one year old the first time that it was sighted. Female humpback whales in the Gulf of Maine have been shown to reach sexual maturity at an average age of five years (Clapham, 1992), a figure that corresponds well with findings for both male and female humpback whales in the Southern Hemisphere (Chittleborough, 1965). Animals known to be less than five years old were considered to be juvenile, while those known to be at least five years old were considered to be sexually mature. A maturational class could not be confidently assigned to whales that were not seen as calves and who were first catalogued less than four years before the study period. For stranded animals, age class was also inferred from body length. Animals less than 9.9m were considered to be dependent or newly independent animals born the previous winter. Males between 10.0 and 11.5m and females between

10.0 and 11.9m were considered to be independent but sexually immature. Males greater than 11.5m and females in excess of 11.9m were considered to be sexually mature (Clapham and Mead, 1999).

The sex of live animals was determined by molecular genetic analysis of skin samples collected by biopsy techniques, either in the USMA or the Gulf of Maine (Bérubé and Palsbøll, 1996a; b; Palsbøll *et al.*, 1997). The sex of dead whales was determined by direct examination of the genital area and/or reproductive organs.

For the purpose of discussing temporal distribution of records, the seasons were defined as follows: spring = April-June; summer = July-September; autumn = October-December; and winter = January-March.

RESULTS

Strandings[1]

Of 52 known humpback whale mortalities in the USMA between 1990 and 2000, 32 were from North Carolina, 13 from Virginia, 3 from Delaware and 2 each from Maryland and New Jersey. These are summarised by year in Fig. 2 and by month in Fig. 3. In addition, one individual originally photographed in the Gulf of Maine in the winter of 2000 stranded in Virginia in April of 2001. This whale was not included in stranding analyses (Table 1), but was included in the matching analyses (Tables 4 and 5).

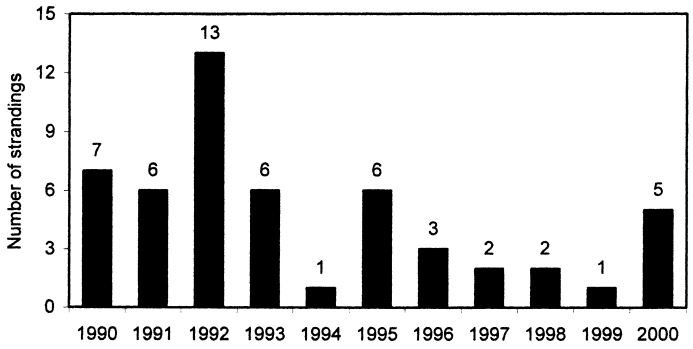


Fig. 2. Yearly frequency of humpback whale strandings in the US mid-Atlantic (NJ-NC) from 1990-2000 ($n = 52$).

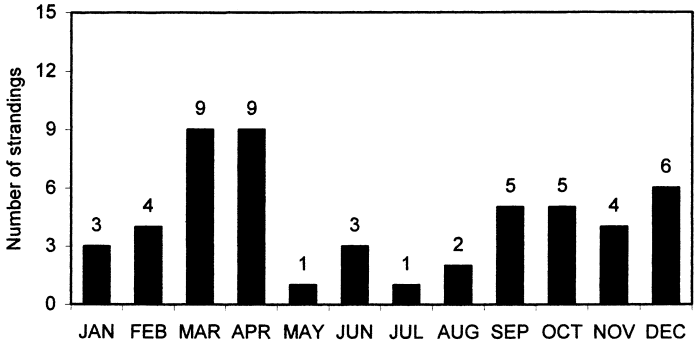


Fig. 3. Monthly frequency of humpback whale strandings in the US mid-Atlantic (NJ-NC) from 1990-2000 ($n = 52$).

Strandings occurred in all years, although there was a noticeable peak of 13 animals in 1992. They were also distributed throughout the year, with a peak in late winter/early spring and the fewest strandings (eight) in summer. In 48 cases, there was information on the sex and/or length of the stranded animal; these are summarised by sex and age class (inferred from length) in Table 1. The sex ratio of stranded whales did not differ significantly from parity ($\chi_2 = 0.20$, df = 1, $p = 0.65$). Judged by length data alone, the majority (39, or 81.2%) would be considered first-year whales, including eight animals that were likely to have been dependent (unweaned) calves. Seven (14.6%) would be considered sexually immature, while only two (4.2%) were large enough to be defined as adults. However, five animals had a prior sighting history from the Gulf of Maine with which to evaluate these length-based determinations. One 7.2m female was first documented as a calf and was therefore known to have died in the spring of her second year. Another male was known to have been at least four years old, despite a measured length of only 9.9m. The exact age of the remaining three animals was unknown, but they were at least near the end of their second year (9.5m male) or the beginning of their third year (8.9m female, 11.1m male) when their carcasses were recovered. Based on carcass length alone, all but the last animal listed would have been considered first-year whales. In addition, necropsy reports of some animals indicated the presence of partially digested fish and/or fish bones and otoliths in the stomachs of stranded whales.

Table 1

Strandings of humpback whales in the US mid-Atlantic, 1990 to 2000, by sex and age class. Age class was inferred from length, as described in Methods. However, some designations are known from sighting history data to be inaccurate (see text).

Age class	Male	Female	Unknown	Total	
1st-year whale	17	16	6	39	81.2%
Immature	4	2	1	7	14.6%
Mature	1	-	1	2	4.2%
Total	22	18	8	48	100.0%

Live whale sightings[1]

Most sightings of live whales were documented in Virginia, which was the focus of most of the sighting effort. However, sightings were also made in New Jersey and North Carolina. Humpback whales were sighted in all seasons, but predominated between January and March. Of the 58 live whales seen, 43 were sighted in winter, primarily in Virginia and North Carolina, 6 were sighted in spring, 12 were observed in summer, primarily off New Jersey and 3 were observed in the autumn. Four whales were sighted in two seasons, one was observed in three seasons. In addition to identifying photographs, whale watch interpreters in Virginia have often noted observing whales with extended ventral grooves surfacing with mouths partially open (Virginia Marine Science Museum, unpublished data).

Thirteen individuals were documented in the USMA region in multiple years. Eleven were sighted in two years, while two were sighted in three years. Most ($n = 8$) re-sightings were limited to consecutive years. However, animals were also found to return 2 ($n = 2$), 3 ($n = 2$), or 6 years ($n = 1$) after their initial sighting.

Exact or minimum ages were available for ten individuals previously documented in other regions. The only live whale first documented as a calf was subsequently sighted in the USMA at the age of seven. The average minimum age of the remaining sample was 3.7 years ($n = 14$, range = 1-6 yrs). When exact and minimum age data were combined, four sightings (26.7%) involved animals that had exceeded the average age of sexual maturity (five years). An additional

[1] The complete set of stranding and sightings data are available from the author or the Office of the Journal.

five sightings involved animals that were at least four years old at the time. As females in the Gulf of Maine have been sighted with calves as early as the age of five, it is conceivable that some of those animals might also have reached sexual maturity.

The sex of 10 USMA whales was determined from molecular genetic analysis of biopsy samples; seven were females and three were males.

Results of photographic matching

A minimum of 44 animals were documented alive in the USMA between 1990 and 2000. Because some individuals were represented only by dorsal fin and others only by fluke photographs, it is likely that some, but not all, animals were catalogued more than once. If however, all of the dorsal fins were different from the flukes, there were as many as 58 unique individuals in the dataset. After rejection of poor quality photographs, 22 unique animals could be identified by fluke pattern. Of the 52 stranded whales, 19 had fluke images of adequate condition for identification.

Sightings information on matched live and dead whales is summarised in Tables 2 and 3, respectively. Three additional Gulf of Maine matches to live whales were made on the basis of dorsal fin photographs; however, these are not included in Tables 2-5 because this feature was not available for matching to catalogues from other feeding grounds. A summary of the number of individuals identified, by state, is given in Table 4. Fluke photographs from 41 individual humpback whales (22 living and 19 dead) were deemed suitable for comparison to other catalogues from the North Atlantic. As shown in Table 5, 20 (14 live and 6 dead) of the 41 individuals were matched to the following areas: Gulf of Maine (16; 39.0%), Newfoundland (5; 12.2%) and the Gulf of St Lawrence (1; 2.4%). Two live whales were matched to both the Gulf of Maine and Newfoundland making a total of 22 matches of 20 individuals. One Gulf of Maine whale was also seen in the West Indies (on Silver Bank) three years prior to its sighting in the USMA.

DISCUSSION

Habitat use and population composition

Swingle *et al.* (1993) presented data on sightings of humpback whales feeding in the nearshore waters of Virginia during the winters of 1991 and 1992. They suggested that the increase in records of this species from the USMA was a recent phenomenon and was not likely due to increased observer coverage. A similar conclusion was reached by Wiley *et al.* (1995), who reviewed stranding records of humpback whales along the USMA coastline. They found records of 30 strandings from 1985-1992, with the highest number in North Carolina and the greatest density of strandings occurring in an area from the Chesapeake Bay to Cape Hatteras. Four of the strandings occurred from 1985-1989 (0.8/year) while the remaining 26 were reported from 1990-1992 (8.7/year). Wiley *et al.* (1995) argued that, since large whale strandings are unlikely to escape public notice, the dramatic increase in such events was real and probably reflected an increase in the use of these waters by the species. They considered the possibility that a winter concentration of humpback whales that had always been present in offshore waters had moved inshore in recent years, but noted that this was not supported by data from systematic aerial and shipboard surveys of the region (CeTAP, 1982).

Table 2

Summary of the sighting histories of live whales matched by fluke to high-latitude areas. HWC = North Atlantic Humpback Whale Catalogue number; GOM = Gulf of Maine; NFD = Newfoundland; GSL = Gulf of St Lawrence; * = exact date unknown.

	Mid-Atlantic US sightings			High Latitude matches		
HWC No.	ID	State(s)	Year(s)	ID	Area	Year(s)
0984	MAMNCLI001	NJ, VA	1990, 92, 93	EL CID	GOM	1990
3792	MAMNCLI008	VA	1992, 93	Y2657	NFD	1993
3791	MAMNCLI014	NC	1993	NONE	NFD	1994
3690	MAMNCLI058	NJ	1992, 93	BORDER	NFD/GOM	1990, 92/91
7044	MAMNCLI015	NJ, NC, VA	1993, 94	AETNA	GSL	1991, 93
8243	MAMNCLI019	NJ, NC	1992, 94	PUMICE	GOM	1997, 98
3715	MAMNCLI026	NC	1994	NONE	NFD	1991
3680	MAMNCLI032	VA	1993, 95	JETTY	NFD/GOM	1990/ 92, 93
8250	MAMNCLI034	NC	1994	CAMELOT	GOM	1995
8111	MAMNCLI038	VA	1995	HAWKSBILL	GOM	1992, 93, 98
8094	MAMNCLI047	VA	Winter 1999*	SWAN	GOM	1992, 97, 98, 99
8244	MAMNCLI044	VA	Winter 1999*	NIKE	GOM	1994, 98, 99
8252	MAMNCLI045	VA	Winter 1999*	DENALI	GOM	2000
8262	MAMNCLI060	NC	2000	CHARYBDIS	GOM	2000

Table 3

Summary of the sighting histories of dead whales matched by fluke to high-latitude areas. GOM = Gulf of Maine. * This GOM whale is the only 2001 stranding included here. It was not included in data presented in Table 1.

	Mid-Atlantic US sightings			High Latitude matches		
HWC No.	Field number	State(s)	Date	ID	Area	Year(s)
830	13-2-90-DC-W	NC	5 Feb. 1990	COBWEB	GOM	1989
832	VMSM 19901013	VA	19 Nov. 1990	BUBBLES	GOM	1990
996	92-MM-AO-MN-05	MD	16 Apr. 1992	CHOPPER	GOM	1991
887	93-MM-AO-MN-02	DE	6 Mar. 1993	THUNDERBIRD	GOM	1990, 92
8285	VMSM 19961010	VA	2 Apr. 1996	DEXTRA	GOM	1995
8264*	VMSM20011038	VA	9 Apr. 2001	INLAND	GOM	2000

Table 4

Summary of live and dead humpback whales individually identified from fluke photographs taken in the US mid-Atlantic, by state. States are listed from North to South. NJ = New Jersey; DE = Delaware, MD = Maryland; VA = Virginia; NC = North Carolina. * There were five cases in which an individual whale was observed in the waters of more than one state, including one animal seen in three states.

Type	*n*	NJ	DE	MD	VA	NC
Live animals*	22	8	0	0	11	9
Dead animals	19	1	1	1	7	9
Total	41	9	1	1	18	18

Table 5

Summary of fluke matches of live (n =14) and dead (n = 6) humpback whales photographed in the US mid-Atlantic. One Gulf of Maine whale was also seen in the West Indies (on Silver Bank) three years prior to its sighting in the mid-Atlantic. Percentages likely represent minimum figures for match rates; see text for discussion of biases. *Two whales were observed in both the Gulf of Maine and in Newfoundland resulting in 16 matches with 14 individuals.

Type	*n*	Gulf of Maine	Newfoundland	Gulf of St Lawrence
Live animals*	22	10 (45.5%)	5 (22.7%)	1 (4.5%)
Dead animals	19	6 (31.6%)	0	0
Total	41	16 (39.0%)	5 (12.2%)	1 (2.4%)

The data presented in this study can add little to the above discussion. Whether the appearance of humpback whales in the USMA coastal waters reflects population expansion in the North Atlantic, notably in the Gulf of Maine and off Atlantic Canada, is unclear. However, given the estimated population growth rate of 6.5% in at least the former region during the 1980s (Barlow and Clapham, 1997), this is not an unreasonable explanation.

Existing data continue to support the hypothesis that humpback whales use the USMA waters primarily during the winter, with some additional occupation at other times of year. The relative paucity of humpback whales in much of the region during summer was confirmed by systematic aerial surveys conducted by the Northeast Fisheries Science Center in 1995 and 1998. These surveys covered the region from Cape Hatteras, North Carolina north to New Jersey and found no humpback whales on the continental shelf (D. Palka and G. Waring, unpublished data).

Both previous studies from this region (Swingle *et al.*, 1993; Wiley *et al.*, 1995) used body length data from strandings as well as visual estimates of length from live whale observations to suggest that local population composition was heavily biased towards sexually immature animals. Of 25 stranded humpbacks from 1985-1992 for which body length was measured, all were below mean lengths for sexual maturity (11.6m for males and 12.0m for females). Seventeen of the 25 were between 8.0 and 9.9m in length and were thus considered by Wiley *et al.* (1995) to be newly independent (i.e. animals at the end of their natal year and recently separated from their mothers); this is in agreement with more recent work defining length at birth and independence in this species (Clapham *et al.*, 1999).

In contrast, the data reported here tell a somewhat different story about the composition of the USMA animals. Assuming that the average age at attainment of sexual maturity is five years (Clapham, 1992), at least 26.7% of the live whales previously catalogued outside of the region were likely to have been reproductively mature when seen in the USMA based upon the length of their sighting histories. It should be noted that this sample may not be representative of the USMA animals since the likelihood of cataloguing an individual increases over time. However, the documentation of calves is not subject to the same bias, and the only live whale first seen at that age was known to be seven years old at the time of its USMA sighting.

It is not known whether the presence of older whales in the sample represents a change in the composition of the population or is an artifact of the different methodologies employed. In the case of free-ranging whales, visual size estimates have not been tested on a sample of known-age whales, and their subjectivity can produce unreliable results, even among experienced observers. It is also possible that the USMA whales may be physically small for their age. In this study, stranded animals with previous sighting histories were found to be older than their carcass lengths suggested. Under-development, particularly for females, may reduce reproductive fitness and might explain the presence of animals of reproductive age outside the breeding range in winter.

Taking all the data together, we suggest that the USMA waters represent a supplemental feeding ground which may be used by juvenile as well as mature humpback whales primarily in winter but also at other times of the year. That some individuals were observed in more than one year indicates repeated use of this habitat. The occurrence in the stranded sample of apparently yearling whales might indicate that the area (or somewhere nearby) constitutes a place where some mothers wean and separate from their calves; however, there is no way to test this hypothesis at present. The presence of other small animals may reflect overwintering by whales that, while not yearlings, are not yet sexually mature and thus have less reason than adult humpback whales to undertake the migration to tropical breeding areas. However, there is little data on residency times of individuals to confirm that overwintering takes place.

The adults observed in the study area might be passing through on their migration to or from the West Indies, or they may be whales that have chosen to remain in higher latitudes rather than migrating south. However, this question cannot be resolved at present. Neither the available individual USMA residency data nor the available high latitude sightings data preclude either possibility.

Population identity

Gulf of Maine

Of the four western North Atlantic feeding stocks (Gulf of Maine, Newfoundland/Labrador, Gulf of St Lawrence and West Greenland), the Gulf of Maine is geographically the closest to the study area and might be expected to provide most or all of the USMA animals.

The observed rates of exchange between the USMA and the Gulf of Maine (45.5% for live whales and 31.6% for stranded animals) are indeed high, exceeding exchange rates documented among the major known North Atlantic feeding grounds by more than an order of magnitude. For example, of 1,082 Gulf of Maine humpbacks compared to other North Atlantic regions through the end of 2000, only 12 (1.1%) were also recorded off Newfoundland and 22 (2.0%) in the Gulf of St Lawrence (J. Allen, unpublished data). Furthermore, the magnitude of exchange reported here should be considered a minimum estimate. USMA individuals cannot be excluded from the Gulf of Maine aggregation just because they were not successfully matched

to the catalogue. Despite the fact that sampling occurs throughout the Gulf of Maine, sightings effort in Massachusetts Bay, in particular, has been substantially higher than in other regions. Individuals that consistently use other areas of the Gulf of Maine would therefore have been less likely to be sighted and catalogued. Gulf of Maine sightings of matched USMA whales occurred primarily within areas of the highest GOM sighting effort, suggesting that effort may have played a role in the matching results. Effort may also explain why the fluke-based match rate was lower for USMA stranded whales than for live whales; because the likelihood of detecting a Gulf of Maine whale increases with time, it is conceivable that fewer were matched because fewer opportunities existed to sample them alive.

Atlantic Canada

In addition to the expected Gulf of Maine matches, it was somewhat surprising to also find several matches to Atlantic Canada (five to Newfoundland and one to the Gulf of St Lawrence) in the sample. Although these six matches represent only 27.3% of the sample of live whales, this greatly exceeds the among-feeding-grounds match rates noted above.

Furthermore, there is good reason to believe that the observed match rate significantly under-represents the true presence of Canadian whales in the study area. Although photo-identification effort has been relatively constant in the Gulf of St Lawrence, such effort in Newfoundland was drastically reduced after 1993 following completion of the YONAH project. Indeed, all five matches to Newfoundland were observed there prior to 1995; of the ten live whales matched to the Gulf of Maine, six (60.0%) had sighting histories which predated 1994.

An additional bias relates to the age of animals found off the USMA. Although it is clear from the sighting histories of some of the known Gulf of Maine whales in the sample, that a range of age classes is represented in the region, the body length data from stranded animals leave little doubt that many of the whales occurring there are juveniles (i.e. younger than five years of age). Given the lack of photo-identification effort in Newfoundland after 1993, only older whales photographed in the USMA would be matchable to that area of Canada. Put another way, there was virtually no sample of Newfoundland whales to match to after 1993, thus any young animals from that area photographed off the USMA in subsequent years would necessarily have been of unknown origin.

CONCLUSION

Whatever the true representation of Canadian whales in the study area, it is apparent that the waters of the USMA are a mixing area for humpback whales from different summer feeding grounds. Although the parallels can be taken only so far, a similar situation appears to exist for harbour porpoises (*Phocoena phocoena*) in this area. Recent genetic studies have suggested that porpoises found off the USMA have more than a single stock origin, with some animals perhaps coming from as far away as Greenland (Rosel *et al.*, 1999). However, since the genetic analysis was based upon the frequency distribution of mitochondrial DNA haplotypes in a relatively small sample (rather than on resightings of identified individual porpoises), this conclusion must remain tentative pending further research.

Management implications

The mixing of humpback whales from different feeding stocks in the study area will mean that the mortalities occurring there will have a lower population impact than if all whales came from a single population. The significance of this will depend on whether whales from each feeding stock are equally likely to die in the study area, a question which cannot be addressed with the existing data. Although the only matches to dead whales were to the Gulf of Maine, the biases noted above —regarding both sampling effort elsewhere and the age of the animals in question —make establishing matches to other regions less likely.

In this context, it is worth noting that a humpback whale that stranded in Brooklyn, New York on 5 September 1991 was photographed alive off Newfoundland the previous year (J. Allen and R. Seton, unpublished data). Since the stranding location is not part of the USMA, this whale was not included in the sample used here. However, New York abuts the northern margin of the study area; thus, the presence of a Canadian whale there lends support to the idea that USMA mortalities may involve more than a single population.

The impact of USMA mortalities on the feeding populations concerned cannot be reliably assessed at present. To do this would require the ability to assign mortalities to the different feeding aggregations and also to have reliable and reasonably precise estimates of abundance for those stocks. There is no current estimate for Newfoundland or the Gulf of St Lawrence, although older data from the YONAH project indicate that at least the former population is 'large' (Smith *et al.*, 1999). The Gulf of Maine population was recently estimated from line transect data at 816 (CV = 0.45, Clapham *et al.*, 2001), not statistically different from the YONAH estimate of 652 (CV = 0.29). Both estimates are subject to possible negative biases (IWC, 2002).

There are no data on population growth rates for humpback whales off Atlantic Canada. The Gulf of Maine population was growing at an estimated rate of 6.5% in the 1980s and early 1990s (Barlow and Clapham, 1997), which is close to the theoretical maximum for this stock. However, an updated recent analysis of photo-id data from 1992-2000 found that population growth had declined during the period 1992-1995; this was linked to a major decrease in calf survival estimates (Clapham *et al.*, 2001). If this decrease was real, it may be partly related to the high mortality among young whales in the USMA. However, as noted by Clapham *et al.* (2001), the decline in calf survival and in apparent population growth rate coincided exactly with a shift in humpback whale distribution away from intensively studied areas; thus, the decline may be at least partly an artifact of this distribution shift. Further analysis of survival and population growth in the Gulf of Maine will require additional years of photo-id data.

In order to determine whether the observed level of mortality in the USMA represents a potentially serious issue for the Gulf of Maine feeding aggregation, further study in the USMA region, including continued monitoring of mortalities, should be a high priority for management.

ACKNOWLEDGEMENTS

We thank Per Palsbøll and Martine Bérubé for molecular genetic sex determination. We are grateful to the contributors to the North Atlantic Humpback Whale Catalogue (particularly Richard Sears of Mingan Island Cetacean Study and Jon Lien, Whale Research Group at

Memorial University of Newfoundland), the Mid-Atlantic Humpback Whale Catalogue and to the many investigators and field staff of the Years of the North Atlantic Humpback Whale Project. Richard Merrick and Fred Serchuk of the Northeast Fisheries Science Center, NOAA Fisheries reviewed an earlier draft of this manuscript. We also wish to express our thanks to members of the Northeast and Southeast Regional Stranding Networks who investigated the stranded specimens. Malcolm Martin played an integral role in collecting data and photographs in the field and Josh Miller (Center for Oceanic Research and Education) provided supplemental sighting history data for one Gulf of Maine whale. This work was funded in part by the US National Marine Fisheries Service, the Virginia Marine Science Museum Foundation and the Whale and Dolphin Conservation Society. We thank two anonymous reviewers for their helpful suggestions on improving the text.

REFERENCES

Barlow, J. and Clapham, P.J. 1997. A new birth-interval approach to estimating demographic parameters of humpback whales. *Ecology* 78(2):535-46.

Bérubé, M. and Palsbøll, P.J. 1996a. Erratum of 'Identification of sex in cetaceans by multiplexing with three ZFX and ZFY specific primers'. *Mol. Ecol.* 5:602.

Bérubé, M. and Palsbøll, P.J. 1996b. Identification of sex in cetaceans by multiplexing three ZFX and ZFY specific primers. *Mol. Ecol.* 5:283-7.

CeTAP. 1982. A characterization of marine mammals and turtles in the mid- and north Atlantic areas of the US outer continental shelf. Cetacean and Turtle Assessment Program, University of Rhode Island. Final Report, Contract AA51-C78-48. Bureau of Land Management, Washington, DC. 538pp. [Available from http://www.blm.gov]

Charif, R.A., Clapham, P.J. and Clark, C.W. 2001. Acoustic detections of singing humpback whales in deep waters off the British Isles. *Mar. Mammal Sci.* 17:751-68.

Chittleborough, R.G. 1965. Dynamics of two populations of the humpback whale, *Megaptera novaeangliae* (Borowski). *Aust. J. Mar. Freshwater Res.* 16:33-128.

Clapham, P.J. 1992. Age at attainment of sexual maturity in humpback whales, *Megaptera novaeangliae*. *Can. J. Zool.* 70(7):1470-2.

Clapham, P.J. and Mayo, C.A. 1987. Reproduction and recruitment of individually identified humpback whales, *Megaptera novaeangliae*, observed in Massachusetts Bay, 1979-1985. *Can. J. Zool.* 65(12):2853-63.

Clapham, P.J. and Mead, J.G. 1999. *Megaptera novaeangliae*. *Mamm. Species* 604:1-9.

Clapham, P.J., Wetmore, S.E., Smith, T.D. and Mead, J.G. 1999. Length at birth and at independence in humpback whales. *J. Cetacean Res. Manage.* 1(2):141-6.

Clapham, P., Barlow, J., Cole, T., Mattila, D., Pace, R., Palka, D. and Robbins, J. 2001. Stock definition, abundance and demographic parameters of humpback whales from the Gulf of Maine. Paper SC/53/NAH10 presented to the IWC Scientific Committee, July 2001, London (unpublished). [Paper available from the Office of this Journal].

International Whaling Commission. 2002. Report of the Scientific Committee. Annex H. Report of the Sub-Committee on the Comprehensive Assessment of North Atlantic Humpback Whales. *J. Cetacean Res. Manage. (Suppl.)* 4:230-60.

Katona, S.K. and Beard, J.A. 1990. Population size, migrations and feeding aggregations of the humpback whale (*Megaptera novaeangliae*) in the western North Atlantic Ocean. *Rep. int. Whal. Commn* (special issue) 12:295-305.

Katona, S.K. and Whitehead, H.P. 1981. Identifying humpback whales using their natural markings. *Polar Rec.* 20:439-44.

Larsen, A.H., Sigurjónsson, J., Øien, N., Víkingsson, G. and Palsbøll, P.J. 1996. Population genetic analysis of mitochondrial and nuclear genetic loci in skin biopsies collected from central and northeastern North Atlantic humpback whales (*Megaptera novaeangliae*): population identity and migratory destinations. *Proc. R. Soc. Lond. Ser. B.* 263:1611-8.

Palsbøll, P.J., Allen, J., Bérubé, M., Clapham, P.J., Feddersen, T.P., Hammond, P.S., Hudson, R.R., Jórgensen, H., Katona, S., Larsen, A.H., Larsen, F., Lien, J., Mattila, D.K., Sigurjónsson, J., Sears, R., Smith, T., Sponer, R., Stevick, P. and Øien, N. 1997. Genetic tagging of humpback whales. *Nature, Lond.* 388:767-9.

Rosel, P.E., France, S.F., Wang, J.Y. and Kocher, T.D. 1999. Genetic structure of harbour porpoise *Phocoena phocoena* populations in the Northwest Atlantic based on mitochondrial and nuclear markers. *Mol. Ecol.* 8:S41-54.

Smith, T.D., Allen, J., Clapham, P.J., Hammond, P.S., Katona, S., Larsen, F., Lien, J., Mattila, D., Palsbøll, P.J., Sigurjónsson, J., Stevick, P.T. and Øien, N. 1999. An ocean-basin-wide mark-recapture study of the North Atlantic humpback whale (*Megaptera novaeangliae*). *Mar. Mammal Sci.* 15(1):1-32.

Swingle, W.M. and Barco, S.G. 2000. Mid-Atlantic humpback whale catalog: live sightings and strandings. VMSM Scientific Report 2000-1 (unpublished). 80pp. [Available from: Virginia Marine Science Museum, Virginia Beach, VA 23451].

Swingle, W.M., Barco, S.G., Pitchford, T.D., McLellan, W.A. and Pabst, D.A. 1993. Appearance of juvenile humpback whales feeding in the nearshore waters of Virginia. *Mar. Mammal Sci.* 9(3):309-15.

Waring, G.T., Quintal, J.M. and Swartz, S.L. (eds.). 2000. U.S. Atlantic and Gulf of Mexico marine mammal stock assessments −2000. NOAA Technical Memorandum NMFS-NE-162, 303p. [Available from http://www.nmfs.gov]

Wiley, D.N., Asmutis, R.A., Pitchford, T.D. and Gannon, D.P. 1995. Stranding and mortality of humpback whales, *Megaptera novaeangliae*, in the mid-Atlantic and southeast United States, 1985-1992. *Fish. Bull.* 93(1):196-205.

Sexual maturation in male bowhead whales (*Balaena mysticetus*) of the Bering-Chukchi-Beaufort Seas stock

T.M. O'HARA*, J.C. GEORGE*, R.J. TARPLEY+, K. BUREK# AND R.S. SUYDAM*

Contact e-mail: todd.ohara@north-slope.org

ABSTRACT

Since the mid-1970s, study has focused on reproductive biology of female bowhead whales, while little has been described for males. This study evaluates testicular morphology (mass and length) and histology in relation to body length to determine the onset of male sexual maturity. Mean testis mass and mean testis length were highly correlated. Body length and mean testis mass were significantly correlated and an inflection of increased testicular mass occurred at approximately 12.5-13.0m suggesting the onset of puberty, and also indicated by histologic findings. Biological variability and the fact that few male animals have been examined within this critical length cohort do not allow determination of the onset of maturity with higher precision. Too few mature males have been landed in spring to make statistical comparisons of testes mass with autumn-landed animals within specific size cohorts. Two large (15.7m and 17.7m) males landed in spring had relatively small inactive testes and were diagnosed as pseudohermaphrodites; body length and mean testis length and seminiferous tubule diameter (STD) were not correlated with the other 'normal' whales. The smallest male confirmed as mature based on the presence of spermatocytes was 12.7m. The largest testes measured (combined mass 203kg) were from a whale landed in autumn. Mean STD for individual whales ranged from 33.3-170.9μ and increased with mean testis weight and whale length. The STD is similar within a testis regardless of region evaluated, with minor variability. Some variation was noted for transverse sections within a cross section for some whales but no pattern was evident.

KEYWORDS: AGE AT SEXUAL MATURITY; ARCTIC; BOWHEAD WHALE; MALE; REPRODUCTION

INTRODUCTION

More effort has been focused on the reproductive biology of female rather than male bowhead whales, *Balaena mysticetus*, of the Bering-Chukchi-Beaufort (BCB) Seas stock (Nerini *et al.*, 1984; Koski *et al.*, 1993; Tarpley *et al.*, 1999a; b; c). Little has been published on male reproduction (Kenny *et al.*, 1981; Medway, 1981; Koski *et al.*, 1993; Tarpley *et al.*, 1995), primarily because an understanding of the reproduction of females is more central to the modelling of populations in the context of various management strategies. However information on male reproduction and breeding behaviour (e.g. seasonality) can improve the models used in management and can provide valuable additional information on the general health of populations (e.g. in the context of *Implementation Reviews* – see IWC, In press). The present study presents information on: testicular morphology; histological evidence of sexual maturation in relation to body length (as a proxy for age); and further description of testis appearance in two animals that were deemed pseudohermaphrodites as previously described by Tarpley *et al.* (1995).

Confirmation of male sexual maturation requires determining the presence of spermatozoa either through seminal smears from fresh epididymides or histological examination of testis and epididymis tissue. When these data are collected in concert with data on testis size (length, width, depth and/or mass) for a large number of animals, it may become possible to predict whether animals are mature or immature from the size of testes alone. Testes of intermediate size will still need to be examined histologically to determine whether an animal is pubertal, or mature with recrudescent testes outside the breeding season (if seasonality occurs). In a synthesis of the existing bowhead whale literature, Koski *et al.* (1993) reported 13.0-13.5m as a mean length at sexual maturity for females and 12.0-13.0m for males, based primarily on males having a smaller body length at physical maturity. This study presents additional information on morphometrics of bowhead whales (i.e. standard length) and of their testes (length and mass); and histological indicators of sexual maturity (presence of sperm and seminiferous tubule diameter between and within testis) to estimate the body length range at sexual maturation.

MATERIALS AND METHODS

Examination and tissue sampling of bowhead whales

A subsistence fishery for bowhead whales is undertaken by Alaskan Eskimos (Inuit) during the spring and autumn migrations. Whales landed at Barrow and occasionally in other villages along the northern coast of Alaska were examined by North Slope Borough (NSB) Department of Wildlife Management (DWM) and/or Alaska Eskimo Whaling Commission (AEWC) personnel. Measurements were taken of standard body length, sex, and up to 43 morphological measurements including testicular mass, greatest testicular circumference, and/or length (pole to pole) for one or both gonads per animal. Where possible, epididymal or testicular fluid was collected. All tissue sections were saved in neutral buffered 10% formalin.

A triple beam balance was used to weigh testis < 3.0kg, a 45.3kg (+/- 0.45kg) spring scale for testis < 40kg, and a 136kg spring scale (+/-2.2kg) for larger testis. A transverse core sample (capsule to centre of cross section) from each pole (anterior and posterior) and mid-section (equatorial) plane (cross section) were taken and fixed in formalin for selected whales (Fig. 1). These three different planes or cross sections were further sub-divided from the capsular surface to the centre along the transverse core (Fig. 1). Transverse cores (or entire cross sections with core removed later) were taken from the posterior and anterior poles, and equatorial

* *Department of Wildlife Management, North Slope Borough, Box 69, Barrow, Alaska, USA.*
+ *Department of Veterinary Anatomy and Public Health, College of Veterinary Medicine, Texas A&M University, College Station, Texas, USA.*
# *Alaska Veterinary Pathology Services, PO Box 773072, Eagle River, Alaska, USA.*

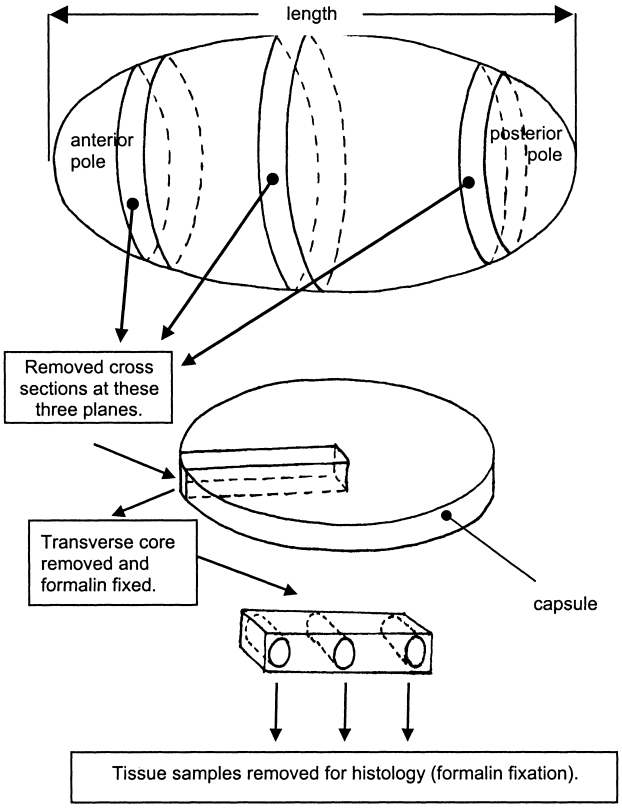


Fig. 1. Diagram of the measuring and sampling sites of a bowhead whale testis.

(mid-section) cross section early in the study (1980-1984) to determine if significant regional differences (between the three planes and within the cross-section or plane) existed. In later samples (1992-1998), only equatorial sections (mid-sections) were taken and sampled near the capsule, in the centre of the cross section, and equidistant between these two landmarks (three tissue samples per testis).

Histological examination

Assessment of maturity

Samples of testicular tissue were trimmed, embedded in paraffin blocks and thin sections were placed on a glass slide and stained with haematoxylin and eosin for microscopic examination following Tarpley *et al.* (1995). Histological examination was similar to that described by Clarke *et al.* (1994) and Tarpley *et al.* (1995) and included: (1) a morphological description; (2) measurement of ten seminiferous tubular cross section diameters (STD) per glass slide or site sampled; and (3) a count of the number of interstitial cells (i.e. Leydig cells) per high dry (40X) field (HDF). For each testis, Clarke (1956) used the mean of 10 tubule diameters; Clarke *et al.* (1994) found no statistical differences between measuring 10 and 20 tubules. For some whales, impression smears and/or wet smears were made from the epididymis and slides were examined under a light microscope (40x). A whale exhibiting 'copious' mature spermatocytes was considered to be sexually mature.

The STD was measured based on the opening in the stroma, as the epithelial cells were often slightly displaced from the supporting stroma and distorted due to processing. Two measurements were taken at right angles to one another across the tubule to determine whether the section was round and therefore a true cross section; only 'round' tubules were measured. Tubules that were closed, with an epithelium with no evidence of mitoses and an approximate tubular diameter of 60μ were considered immature. Tubules that were open with evidence of mitotic activity with a thickened epithelium and a diameter of approximately 130-140μ, were considered mature. Animals with intermediate (approximate diameter 85μ) sized tubules that were open but lacked mitotic activity with a flat epithelium were suspected to be pubertal.

Site of sampling comparison

Formalin-fixed sections of tissue for embedding into paraffin blocks were collected from the transverse core starting inside the capsule (outer connective tissue layer of the testis) and sequentially sampled from this point until reaching the centre of the testis (Fig. 1). These sections were labelled (A, B, C, etc.) sequentially until the centre was reached. For testes collected from 1980-1984, this procedure was conducted for the anterior and posterior poles, and the equatorial cross sections of the testis. For each site sampled (i.e. anterior pole section A) 10 different STDs were measured and the mean STD determined. This was done for 10 male whales landed in the spring, ranging in body length from 7.8-13.6m, and did not include the pseudohermaphrodites. Three different sites of the equatorial (mid-section) plane were examined for samples collected from 1992-1998; the central site, peripheral site and a site equidistant between these two sites (Fig. 1) and a mean STD for that testis is determined from 30 STD measurements (three sites with 10 STD measurements per site).

Statistics

For testes from whales sampled from 1980-1984, each cross section plane (anterior pole, posterior pole or mid-section) had the transverse core sequentially sampled from the periphery (capsule) to the centre and the STD values were compared by ANOVA within each plane (significance determined at a $p < 0.05$). This within-plane comparison was conducted to determine if STD varied significantly by location of the sample (i.e. centre versus closer to capsule) such that sampling site would affect assessment of 'mature' versus 'immature'. The means of STD values within a plane were used to determine an overall mean and SD for each cross section or plane. ANOVA was used to determine if STD differed by plane (anterior pole vs posterior pole vs mid-section) within a testis. For testes for which STD was determined for three cross sections, an overall mean was also calculated. If two testes were assessed for an animal, the average of both testes was calculated to represent the overall animal STD (mean testis STD). Means, SD (standard deviation), ANOVA, Student's *t*-test (Table 1) and figures were derived using *Microsoft Excel 2000* for Windows. Regression analyses were performed on testis morphology using SPSS PC (7.5 for Windows). Models were fitted using the 'curve estimation' function.

RESULTS

Gross morphology

Testes weights were plotted versus whale length (1980-1998 whales) and a steep inflection of increased mean testes mass occurs at 12.5-13.0m body length (Fig. 2). Whale length was linearly related to the *ln* transformed mean testis mass (Table 2). Whale length versus mean testis length was compared using regression analyses with and without the two pseudohermaphrodites and in both cases the relationship was significant and positive. Removing the two male pseudohermaphrodites from the data achieved a better fit (R^2 increased from 0.68 and 0.79, to 0.91 and 0.91 for

Table 1

Mean, standard deviation (SD), and range for body length, mean testis mass and length, and mean seminiferous tubule diameter (STD) for mature and immature (with and without pseudohermaphrodites) whale classes, and mature whale active or inactive spermatogenesis.

Whale class	Whale length (cm)	Mean testis mass (kg)	Mean length of testis (cm)	Mean STD (μ)
All mature				
Mean	1407.6	53.1	92.3	125.8
SD	105.2	28.8	19.2	24.5
No. of whales	18	13	16	15
Range	1270-1660	11.3-101.6	59.0-134	88.0-170.9
All immature*				
Mean	1059.7	4.27	40.4	51.8
SD	245.5	3.32	12.8	9.74
No. of whales	23	9	13	23
Range	760-1770	0.44-9.8	20.4-58	33.3-69.1
All immature except pseudohermaphrodites				
Mean	1000.6	4.27	37.6	51.5
SD	152.7	3.32	11.6	10.1
No. of whales	21	9	11	21
Range	760-1260	0.44-9.8	20.4-58	33.3-69.1
Mature and active				
Mean	1442.5	62.0	101.6	132.5
SD	119.1	25.6	13. 6	25.6
No. of whales	8	7	6	5
Range	1270-1660	20.6-92.1	84-122	102.8-168.8
Mature and inactive				
Mean	1379.6	42.3	86.7	122.5
SD	889.0	30.7	20.4	24.6
No. of whales	10	6	10	10
Range	1290-1528	11.3-101.6	59-134	88.0-170.9
All mature *v.* all immature				
P value**	< 0.0005	< 0.0005	< 0.0005	< 0.0005
Mature-active *v.* mature-inactive				
P value**	0.130	0.124	0.092	0.336

*Includes two pseudohermaphrodites (Tarpley *et al.,* 1995).
** Student's *t*-test.
Note: for mature whales, 'early mature' was diagnosed for four whales. For immature whales, three were classified pubertal. These seven whales have intermediate STDs (mean 85.3μ) and likely represent pubertal whales.

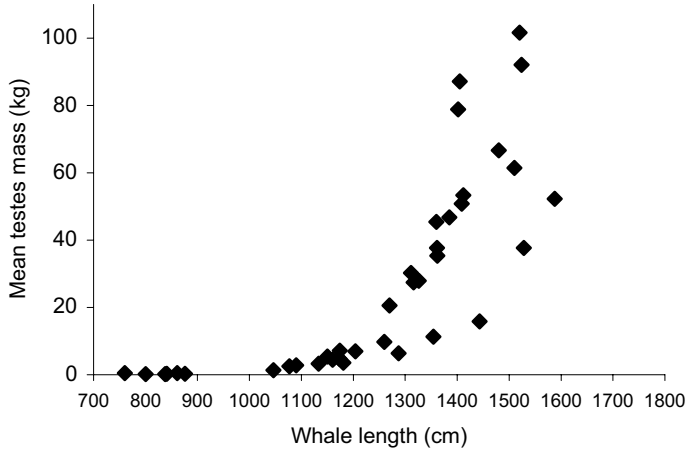


Fig. 2. Mean testis mass (kg) versus whale body length (cm) for select bowhead whales landed in northern Alaska 1980-1998.

non-transformed and *ln* transformed data, respectively) for whale length versus mean testis length (Table 2), and these data are presented in Fig. 3. A strong relationship was found between mean testis mass and mean testis length ($R^2 = 0.89$) and these data are presented in Fig. 4.

Mean testes length increases with body length with a less dramatic inflection (Fig. 3) than seen with mean testis mass, showing more of a linear relationship (Table 2). Two confirmed pseudohermaphrodites fall outside the 'normal' male maturation curve for mean testis length versus body length (Fig. 3); and for mean STD (Fig. 5), as would be expected. Too few mature males have been landed in spring to make statistical comparisons of testicular mass with autumn-landed animals.

Histomorphology

An ANOVA was conducted to determine if a significant difference in STD occurred within each pole and equatorial (mid) cross section for each testis with respect to depth from the capsule along the transverse core sampled. Forty-three separate cross sections (anterior or posterior poles, and equatorial plane) were evaluated in this fashion and 25 cross sections (planes) showed a significant ($p < 0.05$) difference, while 18 did not ($p > 0.05$). However, no pattern was detected and the differences were small in magnitude (approximately 5μ). Ideally, this examination should be conducted on pubertal whale testes.

The average STD for each pole and equatorial cross section was calculated from the averages determined at each depth for testes collected from 1980-1984. An ANOVA compared the means of each pole and equatorial plane for each testis and no significant difference was observed among the two poles and the mid-section for most of the testes evaluated (data not shown). Only two (of 16 tested) testes indicated a regional difference by plane ($P < 0.05$) where the mean STD for one ranged from 36.6-40.1μ, and the other

Table 2

Regression analyses for morphologic measures of testicular mass (kg), length (cm), total whale length (cm), and seminiferous tubule diameters (STD, μ) for male bowhead whales.

Morphology measured	*n*	R^2	*P* value[1] for F statistic	Formula
Testis mass				
Mean testicular mass (x) *v.* mean testicular length	32	0.89	<0.0001	y=27.20+16.91*ln(x)
Whale length (x) *v.* ln mean testis mass [2]	36	0.93	<0.0001	y=-7.75+0.008(x)
Testis length				
Whale length *v.* mean testis length (x) [2]	44	0.68	<0.0001	y=-886.47+133.56*ln(x)
Whale length (x) *v.* ln mean testis length [2]	44	0.79	<0.0001	y=1.243+0.0022(x)
Whale length (x) *v.* mean testis length (w/out pseudos) [3]	42	0.91	<0.0001	$y=4.85*10^{-8}x^{2.93}$
Whale length (x) *v.* ln mean testis length (w/out pseudos) [2,3]	42	0.91	<0.0001	y=0.854+0.0025(x)
Seminiferous tubule diameter (STD)				
Whale length (x) *v.* mean STD	38	0.59	<0.0001	$y=0.00128x^{1.548}$
Mean testis weight (x) *v.* mean STD	17	0.79	<0.0001	y=44.50+21.80*ln(x)
Mean testis length (x) *v.* mean STD	21	0.71	<0.0001	y=-172.65+63.59*ln(x)

[1] *P* value for the ANOVA F-test that coefficient is different from 0.
[2] Data were ln transformed prior to regression analyses.
[3] Pseudohermaphrodites ('pseudos') were determined based on histology and morphology (Tarpley *et al.*, 1995).

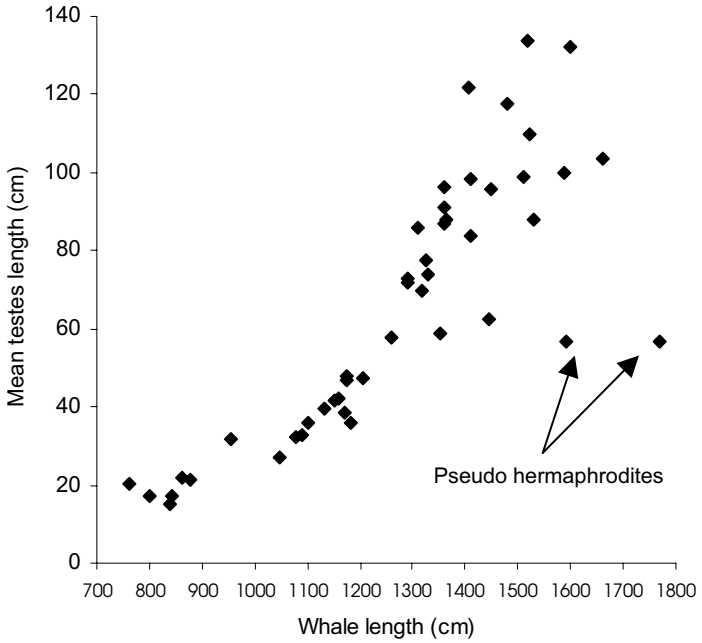


Fig. 3. Mean testis length (cm) versus whale body length (cm) for select bowhead whales landed in northern Alaska 1980-1998.

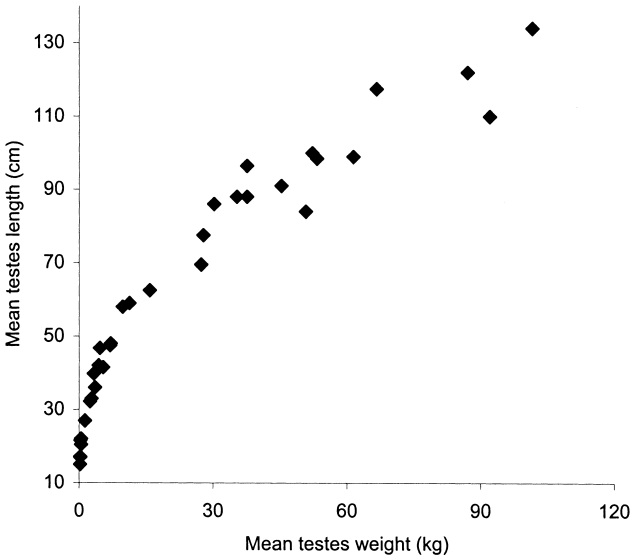


Fig. 4. Mean testis mass (kg) versus mean testis length (cm) for select bowhead whales landed in northern Alaska 1980-1998.

from 41.9-46.5μ. This indicates that mean STD by pole or mid-section does not vary significantly in most specimens; when it did it was only approximately 5μ.

Histological examination indicated that the tubules of immature whales were closed (did not have an obvious lumen), lined by an epithelium characterised by a single layer of prospermatogonia along the basement membrane, Sertoli cells closer to the centre of the tubule, and no evidence of mitoses (Fig. 6). Immature whales (mean whale length of sample 10.6m) had an average STD of 51.8μ (Table 1). Excluding the pseudohermaphroditic whales from the immature category results in a mean STD of 51.5μ (mean whale length of sample 10.0m).

The tubules in mature whales had an obvious lumen, the epithelium was thickened with evidence of mitotic activity, occasional spermatozoa within the lumen, average STD was 125.8μ (Fig. 7) and average whale length was 14.1m (Table 1). Whales with similar findings, but a flat epithelium and no evidence of spermatozoa were classified as suspected mature but inactive. Seven animals had intermediate sized tubules that were open but lacked mitotic activity with a flat

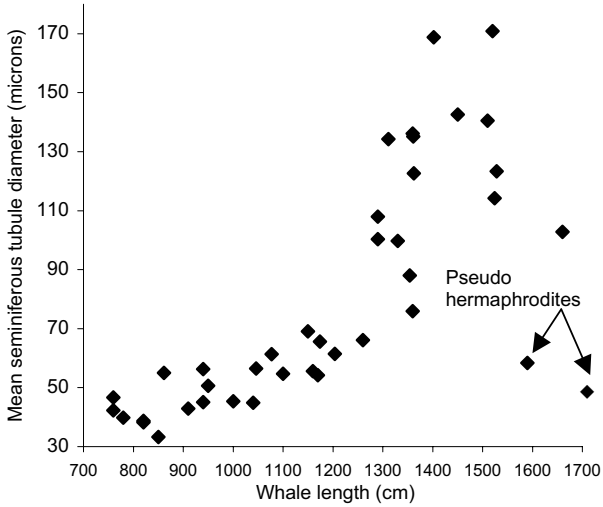


Fig. 5. Mean seminiferous tubule diameter (STD, μμ) versus the body length (cm) for select bowhead whales landed in northern Alaska 1980-1998.

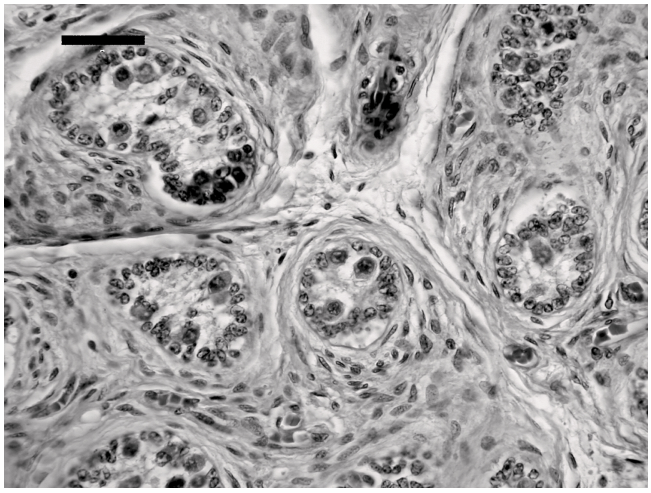


Fig. 6. The seminiferous tubules of immature whales were closed, lined by an epithelium characterised by a single layer of prospermatogonia along the basement membrane, Sertoli cells closer to the centre of the tubule, and no evidence of mitoses. The bar = 40μ.

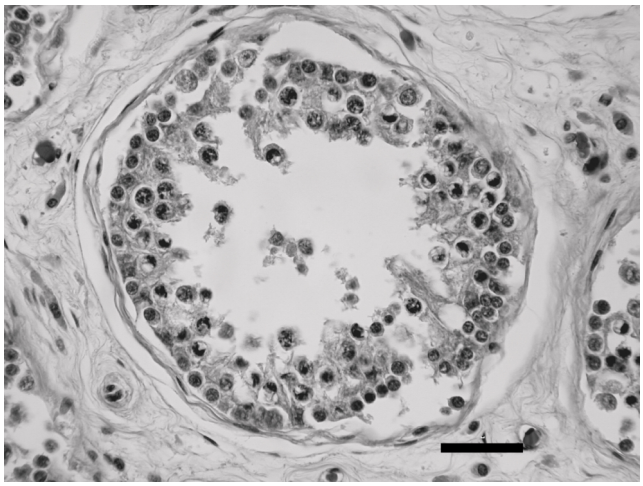


Fig. 7. The seminiferous tubules in mature whales had an obvious lumen and the average STD was 125.8μ. In mature whales with evidence of active spermatogenesis, the epithelium was thickened with evidence of mitotic activity and there were occasional spermatozoa within the lumen. Bar = 40μ.

epithelium and were classified as pubertal or early mature. Pubertal or early mature whales had a mean STD of 85.3μ and mean whale length was 12.6m (range 11.5-13.5m). The smallest whale for which maturity was confirmed was 12.7m in length and harvested in the autumn (Medway, 1981); maturity was determined by the presence of sperm in urine. Whale length, mean testis mass and length, and STD were significantly different for mature versus immature whales (Table 1). None of the endpoints measured comparing histologically determined mature-active to mature-inactive whales proved significantly different (Table 1).

Assessing the number of Leydig cells per HDF was not rewarding and the interstitium varied considerably thus affecting the count. The interstitial cells were very difficult to differentiate. These data are not reported here.

DISCUSSION

Using morphological data (testicular mass and dimensions, whale length) and histological evaluation (STD, presence of spermatozoa and changes in the germinative epithelium), this paper describes the sexual maturation of males of the BCB stock of bowhead whales. One method for estimating the onset of maturity is to determine at what point the testes mass to body length curve inflects sharply. Mean testes mass increases sharply once body length reaches approximately 12.7m, suggesting the onset of puberty can be less than 13.0m. In their review, Koski *et al.* (1993) noted that few data were available on male length at attainment of sexual maturity but predicted that it would occur at 12.0-13.0m. The results here concur with this; 'pubertal' or 'early mature' whales ranged in length from 11.5-13.5m. Durham (1972 *in* Marquette, 1977) reported that sexual maturity in the bowhead whale was estimated to occur at 11.58m but Marquette did not detail how this estimate was obtained. For females, Koski *et al.* (1993) reported values of 14.2m (landed whale data) and 13.0-13.5m (photogrammetric data). However, caution is needed when comparing data from landed whales to photogrammetric studies, as in our experience, stretching of the landed whale will probably increase length measurements. The shorter length at sexual maturation for males is not surprising since females of most baleen whales tend to be larger (Whitehead and Payne, 1981; Koski *et al.*, 1993; George *et al.*, 1999). The estimated age at sexual maturity as determined by stable carbon isotope analysis of the baleen is 17-20yr (Schell *et al.*, 1989; Schell and Saupe, 1993). The estimate mean age for 13m whales using aspartic acid racemisation was about 25 years (George *et al.*, 1999).

The mean STD increased at a body length of approximately 13.0m paralleling an increase in testicular size at a slightly lower body length of about 12.7m. Maturation has been studied in many other whale species (see Boness *et al.*, 2002). Comparisons to these studies of other species indicate male bowhead whales may have the oldest age at sexual maturation (George *et al.*, 1999).

In some testes, a significant difference in STD was detected based on depth from the capsule or plane (anterior pole, posterior pole or equatorial) of sampling, but no clear pattern was evident and the difference was small (approximately 5μ). Thus, we conclude that intratesticular variation in STD does not affect the interpretation of testis maturity or activity if multiple sites are sampled. Whales have been defined as pubertal if immature and mature seminiferous tubules occur in the same testis section (Clarke *et al.*, 1994). In some animals, detection was made of what would appear to be a mean STD that is lower than the large obviously mature males, but higher than clearly immature whales with small STD. These intermediate mean STDs (approximately 85μ) could represent a pubertal or testicular development phase.

Best (1969) showed that sperm whale testes mature from the centre to the periphery and thus location of samples should be considered in any maturation assessment. In our sample, this pattern was only seen in two of 16 whales by comparing cross sections but significant STD differences were found based on depth from the capsule in some sections of 25 of 43 whales. Given this, we recommend sampling from the centre to the periphery of the testis, especially for suspect pubertal animals. With an increased sample of pubertal whales, the centre to periphery maturation theory could be tested in the future. It was reported that on occasion (3-12%), only one testis had detectable sperm in the sperm whale (Clarke *et al.*, 1994) and attempts to evaluate both gonads per bowhead whale should be encouraged.

The largest bowhead whale testes measured in this study had a combined weight of 203.2kg in a 15.2m whale landed in the autumn. Testes of mature northern right whales are often larger, up to 2m long with a combined weight of

1,000kg or more. The relatively large testicular mass for right whales and to a lesser degree bowhead whales may involve a sperm competition reproductive strategy (Brownell and Ralls, 1986). In humpback whales, the paired testes weigh approximately 4.0kg at puberty and are heavier on breeding areas than in feeding areas (Nishiwaki, 1959; Chittleborough, 1965). The combined maximum weight of both testes in the gray whale at the height of the breeding season was approximately 68kg (Rice and Wolman, 1971). The much heavier testes and larger seminiferous tubules of gray whales during the southward migration compared to the northward migration and in the summer feeding areas indicate a marked seasonal sexual cycle with a peak in spermatogenic activity in late autumn and early winter (Rice and Wolman, 1971; Wolman, 1985). Sample size precluded comparisons of spring versus autumn-landed mature males here; bowhead whales are thought to breed in March (Koski *et al.*, 1993).

In conclusion, based on testicular size (mass) and STD, male bowhead whales initiate significant testicular development at approximately 12.5-13.0m in length. In this study, seasonality could not be determined and Leydig cell (interstitial) assessment was found to be of very limited value given their cryptic nature. Future studies utilising special techniques to highlight interstitial cells (i.e. Leydig) and androgen measures in serum and other matrices may help to describe seasonally based changes in testicular activity and better define sexual maturity.

ACKNOWLEDGEMENTS

We greatly appreciate the cooperation of the many whaling captains and crews (Barrow Whaling Captains Association), and the Alaska Eskimo Whaling Commission (AEWC) for allowing us to sample landed whales for this study. We thank L.M. Philo, P. Nader, D. Ramey, V. Woshner, and many others for assistance with sample and data collection. We thank T. Albert, Charles D.N. Brower, and B. Nageak for supporting this study. We thank two anonymous reviewers for their constructive comments. C. Willetto prepared Fig. 1.

REFERENCES

Best, P.B. 1969. The sperm whale (*Physeter catodon*) off the West Coast of South Africa. 3. Reproduction in the male. *Investl Rep. Div. Sea Fish. S. Afr.* 72(3):1-20.

Boness, D.J., Clapham, P.J. and Mesnick, S.L. 2002. Life history and reproductive strategies. pp. 1-432. *In:* A.R. Hoelzel (ed.) *Marine Mammal Biology, An Evolutionary Approach.* Blackwell Publishing, Malden, USA.

Brownell, R.L. and Ralls, K. 1986. Potential for sperm competition in baleen whales. *Rep. int. Whal. Commn* (special issue) 8:97-112.

Chittleborough, R.G. 1965. Dynamics of two populations of the humpback whale, *Megaptera novaeangliae* (Borowski). *Aust. J. Mar. Freshwater Res.* 16:33-128.

Clarke, R. 1956. Sperm whales of the Azores. *Discovery Rep.* 28:237-98.

Clarke, R., Paliza, O. and Aguayo, A.L. 1994. Sperm whales of the southeast Pacific. Part VI. Growth and breeding in the male. *Invest. Cetacea* XXV:93-224.

George, J.C., Bada, J., Zeh, J., Scott, L., Brown, S.E., O'Hara, T. and Suydam, R. 1999. Age and growth estimates of bowhead whales (*Balaena mysticetus*) via aspartic racemization. *Can. J. Zool.* 77:571-80.

International Whaling Commission. In press. Report of the Scientific Committee. *J. Cetacean Res. Manage. (Suppl.)* 5.

Kenny, R.M., Garcia, M.C. and Everitt, J.I. 1981. The biology of the reproductive and endocrine systems of the bowhead whale, *Balaena mysticetus*, as determined by evaluation of the tissues and fluids from subsistence harvested whales. pp. 89-157. *In:* T.F. Albert (ed.) *Tissue Structural Studies and Other Investigations on the Biology of Endangered Whales in the Beaufort Sea*. Report to the Bureau of Land Management from the Department of Veterinary Sciences, University of Maryland, College Park, MD 20742. 953pp. NTIS No. PB86-153566 [Available from: http://www.ntis.gov].

Koski, W.R., Davis, R.A., Miller, G.W. and Withrow, D.E. 1993. Reproduction. pp. 239-74. *In:* J.J. Burns, J.J. Montague and C.J. Cowles (eds.) *The Bowhead Whale*. 1st. Edn. Society for Marine Mammalogy, Lawrence, Kansas. Special Publication. No.2. 787pp.

Marquette, W.M. 1977. *The 1976 Catch of Bowhead Whales* (Balaena mysticetus) *by Alaskan Eskimos with a Review of the Fishery, 1973−1976, and a Biological Summary of the Species*. Northwest and Alaska Fisheries Center, Seattle. 93pp.

Medway, W. 1981. The cytological and clinical evaluation of blood and urine of the bowhead whale, *Balaena mysticetus*. pp. 201-12. *In:* T.F. Albert (ed.) *Tissue Structural Studies and Other Investigations of the Biology of Endangered Whales in the Beaufort Sea*. Report to the Bureau of Land Management from the Department of Veterinary Science, University of Maryland, College Park, MD 20742. 953pp. NTIS No. PB86-153566 [Available from: http://www.ntis.gov].

Nerini, M.K., Braham, H.W., Marquette, W.M. and Rugh, D.J. 1984. Life history of the bowhead whale (Mammalia, Cetacea). *J. Zool., London.* 204:443-68.

Nishiwaki, M. 1959. Humpback whales in Ryukyuan waters. *Sci. Rep. Whales Res. Inst., Tokyo* 14:49-87.

Rice, D.W. and Wolman, A.A. 1971. *The Life History and Ecology of the Gray Whale* (Eschrichtius robustus). American Society of Mammalogists, Special Publication No. 3, Stillwater, Oklahoma. viii+142pp.

Schell, D.M. and Saupe, S.M. 1993. Feeding and growth as indicated by stable isotopes. pp. 491-509. *In:* J.J. Burns, J.J. Montague and C.J. Cowles (eds.) *Special Publication*. No. 2. *The Bowhead Whale*. 1st. Edn. Society of Marine Mammalogy, Lawrence, KS.

Schell, D.M., Saupe, S.M. and Haubenstock, N. 1989. Bowhead (*Balaena mysticetus*) growth and feeding as estimated by d13C techniques. *Mar. Biol.* 103(4):433-43.

Tarpley, R.J., Jarrell, G.H., George, J.C., Cubbage, J. and Stott, G.G. 1995. Male pseudohermaphroditism in the bowhead whale, *Balaena mysticetus*. *J. Mammal.* 76(4):1267-75.

Tarpley, R.J., Hillmann, D.J. and George, J.C. 1999a. Accumulation and persistence of *corpora albicantia* in the bowhead whale ovary. Paper SC/51/AS6 presented to the IWC Scientific Committee, May 1999, Grenada, WI (unpublished). 15pp. [Paper available from the Office of this Journal].

Tarpley, R.J., Hillmann, D.J. and George, J.C. 1999b. Observation on reproductive tract morphology in an immature female bowhead whale. Paper SC/51/AS8 presented to the IWC Scientific Committee, May 1999, Grenada, WI (unpublished). 23pp. [Paper available from the Office of this Journal].

Tarpley, R.J., Hillmann, D.J. and George, J.C. 1999c. Observations on ovarian morphometrics and morphology in the bowhead whale. Paper SC/51/AS7 presented to the IWC Scientific Committee, May 1999, Grenada, WI (unpublished). 45pp. [Paper available from the Office of this Journal].

Whitehead, H. and Payne, R. 1981. New techniques for measuring whales from the air. NTIS No. PB81-161143. US Marine Mammal Commission Report No. MMC-76/22, Washington DC. 36pp. [Available from: http://www.ntis.gov].

Wolman, A.A. 1985. Gray whale, *Eschrichtius robustus* (Lilljeborg, 1861). pp. 67-90. *In:* S.H. Ridgway and R. Harrison (eds.) *Handbook of Marine Mammals*. Vol. 3. *The Sirenians and Baleen Whales*. Academic Press, London and Orlando. xvii+362pp.

A note on the possibility of identifying Leydig and Sertoli cells by immunohistochemistry in bowhead whales (*Balaena mysticetus*)

DEBRA L. MILLER[*], GREGORY D. BOSSART[+], MEHRDAD NADJI[#], RAYMOND TARPLEY[**], BRENDA ROBERTS[++] AND TODD M. O'HARA[°]

Contact e-mail: Todd.O'Hara@north-slope.org

ABSTRACT

Leydig cells have been found to be either unidentifiable or at apparent low numbers during routine histologic examination of bowhead whale testes. Therefore, formalin-fixed, paraffin-embedded testicular tissues from 14 bowhead whales were retrospectively examined to determine if immunohistochemical staining could aid in identification of Sertoli and Leydig cells. Multiple intratesticular samples were examined when available. Sertoli and Leydig cells were differentiated using inhibin and calretinin stains. Significant whale length and seasonal differences were not found; however, a trend toward increased staining intensity was noted for autumn harvested whales.

KEYWORDS: BOWHEAD WHALE; HISTOLOGY; IMMUNOHISTOCHEMISTRY; REPRODUCTION

INTRODUCTION

Obtaining a better understanding of the reproductive process is of particular importance in endangered species management. Such studies are especially challenging in migratory species because indices of reproduction may be affected by the intense physical demands of migration. Additionally, migration patterns often coincide with the breeding season, as is the case for bowhead whales (*Balaena mysticetus*) in Alaska (Koski *et al.*, 1993). The traditional timing of bowhead whale subsistence hunts in Barrow, during the spring and autumn, affords an opportunity for sampling and evaluating reproductive parameters at two periods in the annual migratory cycle (O'Hara *et al.*, 1999).

Biologists are interested in physical indicators of reproductive success. In testes, Sertoli cells are sustentacular or nurse cells for the developing gametes that are positioned basally within mammalian seminiferous tubules (Banks, 1986). They provide the physical support upon (or within) which the spermatogonia are embedded (Banks, 1986). They also secrete oestrogen and are metabolic regulators for developing gametes (Faulkner, 1969). Thus, they participate actively in the movement, development and release of the gametes. This participation makes them potentially valuable reproductive indicators. Similarly, Leydig cells are also a reproductive indicator because of their role in secreting testosterone, the male sex hormone. Leydig cells are the interstitial cells of the testes and are located within the septal connective tissue (Banks, 1986). Testosterone is the key stimulus for development of secondary sex characteristics and reproductive prowess (Faulkner, 1969).

Morphologic parameters (gross and microscopic) of testes have been used in many species to aid in assessment of reproductive status (e.g. Kobayashi *et al.*, 1996; Chapin, 1997; Hopkins and Spitzer, 1997; Short, 1997; O'Hara *et al.*, 2002). For example, in immature animals and during the non-breeding season in seasonal breeders, the epithelium of the seminiferous tubules is quiescent, with only scattered spermatogonia and Sertoli cells present (Banks, 1986). Examination of testicular morphology in bowheads has shown Leydig cells to be either unidentifiable or in low numbers in animals expected to be mature and active (O'Hara *et al.*, 2002). This inability to identify Leydig cells in bowhead whales compromises the assessment of sexual activity and seasonality, which are currently not well described. In humans, differences in the quantity and quality of Sertoli and Leydig cells have been demonstrated by immunohistochemistry (Forti *et al.*, 1992).

Although oestrogen and testosterone markers have been used to identify Sertoli and Leydig cells, respectively, additional markers have been identified and provide a means of identifying non-functional or inactive cells. For Sertoli cells, testis-specific proteins (i.e. inhibin and androgen binding protein) may serve to identify their presence (Forti *et al.*, 1989; Bergh and Cajander, 1990; Vleigen *et al.*, 1993). Specifically, inhibin is a peptide hormone that is a product of primarily Sertoli cells and minimally Leydig cells (Bergh and Cajander, 1990). Likewise, calretinin, a calcium-binding protein important in neurogenesis and neuroprotective mechanisms is thought to correlate with androgen status in males and has been valuable in identifying Leydig cells (Lephart *et al.*, 1998; Lephart and Watson, 1999). The purpose of this study is to identify a technique to specifically distinguish and describe the Leydig and Sertoli cells in

[*] *Department of Pathology, The University of Georgia College of Veterinary Medicine, Veterinary Diagnostic and Investigational Laboratory, Tifton, Georgia, 31793, USA.*
[+] *Harbor Branch Oceanographic Institute, 5600 US 1 North, Fort Pierce, Florida, 34946, USA.*
[#] *Department of Pathology, University of Miami School of Medicine, Miami, Florida, 33136, USA.*
[**] *Texas A & M University, College of Veterinary Medicine, College Station, TX, 77843, USA.*
[++] *Department of Pathology, University of Miami School of Medicine, Miami, Florida, 33136, USA.*
[¥] *North Slope Borough Department of Wildlife Management, Barrow, Alaska, 99723, USA.*

bowhead whale testes. This information will be used to consider seasonality and interpret testosterone data and unusual cases, such as pseudohermaphrodites, in animals harvested in the autumn and spring.

MATERIALS AND METHODS

Sample collection

Testes were opportunistically collected from bowhead whales by the Department of Wildlife Management, North Slope Borough, Alaska. Historically, testicular samples have been collected from three locations, outer (near the tunic), middle and central (O'Hara *et al.*, 2002). All specimens were formalin-fixed, paraffin embedded and prepared on glass slides for staining and histologic examination. Standard haematoxylin and eosin staining protocol was used for initial light microscopic examination.

Immunohistochemistry

Immunohistochemical staining success can be limited by antibody specificity for the targeted antigen. Therefore, species variation can result in false negatives. Lack of specific antibodies for whale tissue antigens necessitates that a chosen target (protein, cell type, product), and associated antibody, is fairly well conserved among mammalian species. This study included testosterone (rabbit anti-testosterone[1]), oestrogen receptor, progesterone receptor and androgen binding protein (mouse anti-human[2]) antibodies for staining testicular tissue. All four of these resulted in negative staining, most likely due to species variation of the antibodies for these products. Antibodies to inhibin (mouse anti-human[3]) and calretinin (rabbit anti-calretinin[4]) yielded positive results for both Sertoli and Leydig cells; however, preferential staining was noted for Sertoli cells by inhibin and Leydig cells by calretinin. Therefore, inhibin and calretinin were used in this study to identify Sertoli and Leydig cells, respectively.

Immunohistochemical staining was performed by the University of Miami Pathology Reference Service Immunoperoxidase Laboratory, Miami, Florida (USA). A labelled streptavidin biotin procedure was used (Nadji and Morales, 1983). Biotin (Vitamin B6) has high affinity binding to avidin (egg white protein), which allows these molecules to be used for 'tagging' different proteins and other molecules. The reaction between biotin and avidin can serve as a low-energy, high-efficiency method of binding substances that are otherwise difficult to bind. In brief, tissue sections approximately 3μm thick were cut from paraffin blocks and placed on glass slides. These slide preparations were deparaffinised, rehydrated and treated with H_2O_2-methanol to block interference from endogenous enzymes. The specific primary antibody was added to the preparation and conjugated with biotin. Next, the detection molecule (peroxidase) congugated to avidin (strepavidin peroxidase) was added and this preparation was incubated at room temperature for 30 minutes. Specific binding of biotin to avidin allows for localisation of the detection molecule over a target antigen, which can then be visualised by addition of a chromagenic substance for the detection molecule (enzyme). Slide preparations were viewed under light microscopy. Staining intensity was graded on a scale of 0-4, with 0 being no staining and 4 being the highest intensity of staining observed.

Morphological measurements

Seminiferous tubules were measured with light microscopy as described in O'Hara *et al.* (2002) from haematoxylin and eosin stained slide preparations. Leydig cells were measured with light microscopy from immunohistochemically stained slide preparations following Clarke *et al.* (1994). In summary, two cells were measured from each of five fields for a total of 10 cells per sample. The widths and the lengths of the Leydig cells (given that they are round to oval structures) were measured using a micrometer and the averages (of the widths and the lengths) were calculated.

Statistics

Means, standard deviation and Student's *t*-test were derived using *Microsoft Excel 2000*. Means and standard deviations were calculated for each dataset (calretinin staining, inhibin staining, seminiferous tubule diameter and Leydig cell dimensions) and within each dataset by season. A Student's *t*-test was used to compare datasets by whale length and by season.

RESULTS

Positive intracytoplasmic staining was present for inhibin and calretinin (Figs 1 and 2). Inhibin staining was generally more intense than the calretinin staining (Figs 1 and 2); although in two pseudohermaphrodites, the staining intensity was reversed. No significant differences ($p > 0.05$) were found in the size of Leydig cells or staining intensity of calretinin or inhibin in relation to whale length (Table 1). The specimens examined were from various intratesticular locations. In one whale, multiple locations were examined but no differences in staining intensity were noted. Significant ($p < 0.05$) differences were noted in seminiferous tubule diameter in comparisons between whales greater and less than 12m, but not in comparisons between whales greater and less than 9m. Different Leydig cell diameters were noted between whales greater and less than 9m if $p < 0.10$ is considered significant, but not between whales greater and less than 12m ($p = 0.16$). Spermatozoa were not observed within the sections used for these immunohistochemical stains.

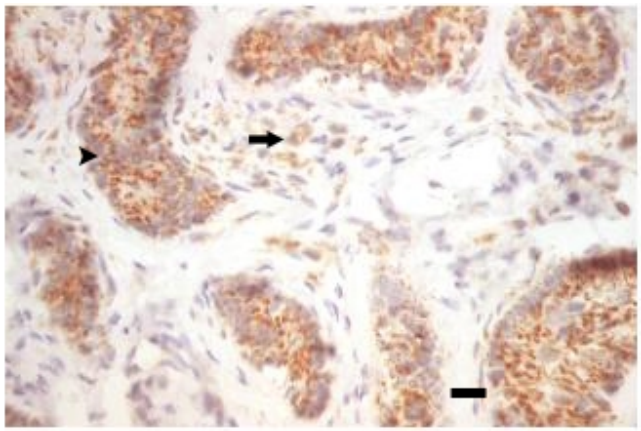


Fig. 1. Positive immunohistochemical staining for inhibin with identification of Sertoli cells (arrowhead) and Leydig cells (arrow) in the testis of a bowhead whale (*Balaena mysticetus*). Bar = 15μ.

[1] BioGenex, 4600 Norris Canyon Road, San Ramon, California, USA.
[2] DAKO Corporation, 6392 Via Real, Carpinteria, California, USA.
[3] Serotec Incorporated, 3200 Atlantic Avenue, Suite 105, Raleigh, North Carolina, USA.
[4] Zymed Laboratories, Inc., 561 Eccles Avenue, South San Francisco, California, USA.

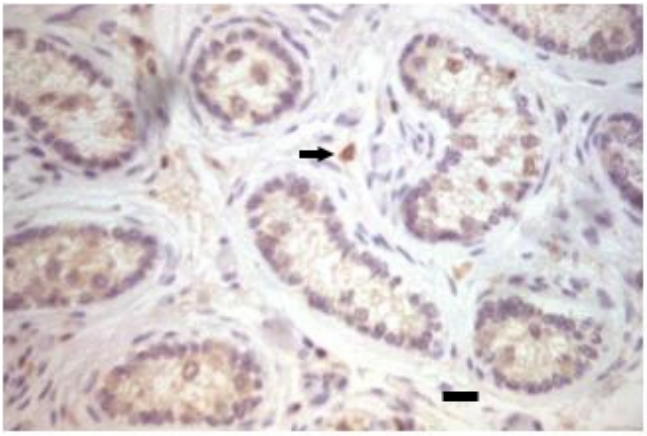


Fig. 2. Positive immunohistochemical staining for calretinin with identification of Leydig cells (arrow) in the testis of a bowhead whale (*Balaena mysticetus*). Bar = 20μ.

DISCUSSION

In general, staining intensity was greater for inhibin than for calretinin; however, Leydig cells were more easily viewed in calretinin stained sections. The variation in intensity might be ascribed to species compatibility to the antibodies and/or quantity of product present in the tissue; however, two pseudohermaphrodites that were included in the study displayed the opposite results for staining intensity (i.e. the calretinin stained much more intensely than the inhibin). Therefore, the quantity of antigen that is present in the tissue is a more likely explanation for this variation in stain intensity than species compatibility.

Sertoli cells are the primary producers of inhibin in males. Leydig cells have been shown to produce minimal amounts of inhibin in humans (Bergh and Cajander, 1990). There remains some debate as to whether or not other testicular sources of inhibin exist but attempts to identify other sources have been inconclusive, especially in adult mammals (Bergh and Cajander, 1990; Meachem *et al.*, 2001). Germ cells have been shown to affect the amount of inhibin production in other mammals but the effects have varied (Bergh and Cajander, 1990; Meachem *et al.*, 2001; Clifton *et al.*, 2002). Additionally, Bergh and Cajander (1990) suggest that inhibin production may reflect both Sertoli and Leydig cell function. The intratubular positive staining for inhibin in the bowhead whale is presumed to reflect the framework provided by the cytoplasm of the Sertoli cells. Other intratubular sources of positive staining were not identified but cannot be ruled out. It is uncertain if the quantity of positive staining for inhibin in the bowhead whale is reflective of either Sertoli or Leydig cell function, but combining these data with serum hormone analysis (i.e. for FSH, inhibin, oestrogen and testosterone) may aid in understanding of the reproductive physiology in this species.

Minimal variation in calretinin and inhibin staining of testicular tissue was noted by length and season. This lack of significant variation in staining intensity by whale length was surprising. Sexual maturity is thought to take place at approximately 25 years of age, which for males corresponds to a length of 12-13m (George *et al.*, 1999) or 12.5-13m (O'Hara *et al.*, 2002). Therefore, males less than 9m are considered immature or young adults. Further investigation, including serum hormone analysis, is required to interpret the positive staining in the immature animals.

Staining intensity seemed to vary between animals that were of similar size but harvested in different seasons (83WW2 versus 82KK1; 96V148 versus 85B2; and 80B5 versus 81H4 versus 98B19). More intense staining was noted among autumn-harvested animals that might indicate seasonal fluctuation in function or differences in sampling or preservation (Table 1). This may suggest a seasonally dependent proliferation or 'turning on' of the cells in response to increased hormonal production, and additional samples are being collected to clarify this. Increased hormonal production in the autumn would coincide with the time preceding migration to the presumed breeding grounds. Koski *et al.* (1993) state that the principal mating period of

Table 1

Immunohistochemical staining intensities (0-4) for calretinin (calret) and inhibin in bowhead whale (*Balaena mysticetus*) testes.

Whale ID	Whale length (m)	Year	Season	Calret	Inhibin	Seminiferous tubule diameter (μm)	Leydig cell dimensions (μm^2)
81WW3 [a]	17.70	1983	S	2	1	50.8	6.2 × 9.2
83WW2	16.20	1983	S	2	2	102.8	9.2 × 10.8
82KK1	16.00	1982	F	2	3	62.0	7.7 × 7.7
83WW1 [a]	15.90	1983	S	3	2	58.4	9.2 × 13.9
84WW2	14.80	1984	S	0	0	75.9	-
96V150	13.60	1996	F	0	0	69.0	-
96V148	12.60	1996	F	4	4	45.8	8.5 × 10.8
85B2	12.20	1985	S	3	4	39.0	9.2 × 13.9
80B5	10.40	1980	S	1	2	44.9	11.6 × 13.9
81H4	10.00	1981	S	1	1	45.4	7.7 × 7.7
98B19	9.52	1998	F	3	4	50.55	8.5 × 8.5
82G1	7.92	1982	S	2	4	37.5	7.7 × 9.2
80WW1	7.80	1980	S	1	2	39.8	7.7 × 7.7
84WW1	7.60	1984	S	2	4	42.3	7.7 × 7.7
Total average	12.30			1.86	2.36	54.58	8.41 × 10.08
Spring average (n=10)	12.05			1.70	2.20	53.68	8.47 × 10.44
Fall average (n=10)	12.93			2.25	2.75	56.85	8.23 × 9.00
Total SD	3.48			1.17	1.50	18.07	1.33 | 2.55
Spring SD	3.85			0.95	1.40	20.77	1.54 | 2.77
Fall SD	2.68			1.71	2.05	10.57	0.46 | 1.61
Students t-test[b] by whale length (<12 *vs* >12)				0.83	0.49	0.04	0.16
Students t-test[b] by whale length (<9 *vs* >9)				0.90	0.31	0.14	0.08
Students t-test[d] by season				0.45	0.56		

[a] Pseudohermaphrodite (PH). [b] Analysis excludes PH.

the bowhead whale remains unknown; however, based on data from foetuses found in females at harvesting, most conceptions may occur during the late winter or spring. Similarly, Mogoe *et al.* (2000) reported a functional reduction of the southern minke whale (*Balaenoptera acutorostrata*) testis during the feeding season; and also found a corresponding decline in plasma testosterone concentration. Contemporaneous plasma testosterone concentrations are needed to further characterise the observations in this bowhead whale study.

A significant difference in seminiferous tubule diameter was observed between whales greater than and less than 12m (approximately the length at sexual maturity). Similar and expected findings were not noted in comparisons of Leydig cell diameters by whale length; possible differences in Leydig cell diameters were noted when comparing whales greater versus less than 9m ($p = 0.08$). Based on their body lengths, these whales were most likely pubertal males and therefore, their testes were undergoing development. Clarke *et al.* (1994) found that Leydig cell size varies with maturity and that pubertal whales have Leydig cells that range from immature to mature depending upon the location within the testes. Adequate intratesticular comparisons regarding Leydig cell diameter cannot be made based on the sections examined in this study.

Negative staining results for animals 84WW2 and 96V150 were most likely due to sample handling. Factors such as tissue autolysis or prolonged formalin storage prior to paraffin embedding can affect immunohistochemical staining (Nadji and Morales, 1983). Times associated with sample handling were not known for these tissues.

The positive staining of the pseudohermaphrodite testicular tissues is interesting and: (1) may indicate testicular function in these animals; or (2) may imply that positive staining in normal tissues cannot be correlated with degree of function or activity. Tarpley *et al.* (1995) noted that the external phenotype in these whales was female but they described the gonads as underdeveloped testes. They noted that the cellular composition of the seminiferous tubules were limited to Sertoli cells and spermatogonia whereas the interstitium was primarily dense collagen, similar to that described for normal males. Active spermatogenesis was not found in the pseudohermaphrodites and anatomically the epididymis, ductus deferens and prostate gland were not present; however, neither were uterine structures. The karyotypes of these whales were XY and chromosomes were morphologically normal with a diploid number of 42. Given these morphological findings, the positive staining for Sertolli cells is not surprising; however, the positive staining for Leydig cells suggests that they may have been present but were not distinguishable with conventional stains (i.e. hematoxylin and eosin) as was the case for the normal males. These data show that such pseudohermaphroditic whales are reproductively deficient based on the anatomical and physiologic studies.

Future work should include correlation of these immunohistochemical findings with oestrogen and testosterone levels from archived serum samples. Correlating these findings with associated serum testosterone and oestrogen levels may aid in interpretation of the reproductive cycle (seasonality) of bowhead whales. Further, unusual cases, such as the pseudohermaphrodites, may be more thoroughly interpreted. These data will improve understanding of the reproductive physiology of this species, particularly the possibility that male bowhead whales vary seasonally in functional maturity and sperm production.

ACKNOWLEDGEMENTS

We thank T. Albert for his support of this project, and for his and V. Woshner's editorial comments in preparation of this manuscript. We also thank the subsistence whale hunters for allowing us to collect the samples used in this study. Many assisted with sample collection and we are grateful for their help, particularly, L.M. Philo, P. Nader, D. Ramey and V. Woshner. Much appreciation goes to M. Dossey for the long hours spent sifting through archived specimens and for packaging and shipping them at a moments notice. Finally, we appreciate the comments of two anonymous reviewers.

REFERENCES

Banks, W.J. 1986. *Applied Veterinary Histology*. 2nd Edn. Williams and Wilkins, Baltimore, MD. 583pp.

Bergh, A. and Cajander, S. 1990. Immunohistochemical localization of inhibin-α in the testes of normal men and in men with testicular disorders. *Int. J. Androl.* 13:463-9.

Chapin, R.E. 1997. The relationships among reproductive endpoints in Swiss mice, using the reproductive assessment by continuous breeding database. *Fundam. Appl. Toxicol.* 38(2):129-42.

Clarke, R., Paliza, O. and Aguayo, A.L. 1994. Sperm whales of the southeast Pacific. Part VI. Growth and breeding in the male. *Invest. Cetacea* XXV:93-224.

Clifton, R.J., O'Donnell, L. and Robertson, D.M. 2002. Pachytene spermatocytes in co-culture inhibit rat Sertoli cell synthesis of inhibin betaB-subunit and inhibin B but not the inhibin alpha-subunit. *J. Endocrinol.* 172(3):565-74.

Faulkner, L.C. 1969. Male reproduction. pp. 155-295. *In:* L.E. McDonald (ed.) *Veterinary Endocrinology and Reproduction*. Lea and Febiger, Philadelphia. 460pp.

Forti, G., Barni, T., Vanelli, B.G., Balboni, G.C., Orlando, C. and Serio, M. 1989. Sertoli cell proteins in the human seminiferous tubule. *Journal of Steroid Biochemistry* 32(18):135-44.

Forti, G., Vannelli, G.B., Barni, T., Balboni, G.C., Orlando, C. and Serio, M. 1992. Sertoli-germ cells interactions in the human testis. *J. Steroid. Biochem. Mol. Biol.* 43(5):419-22.

George, J.C., Bada, J., Zeh, J., Scott, L., Brown, S.E., O'Hara, T. and Suydam, R. 1999. Age and growth estimates of bowhead whales (*Balaena mysticetus*) via aspartic racemization. *Can. J. Zool.* 77:571-80.

Hopkins, F.M. and Spitzer, J.C. 1997. The new Society for Theriogenology breeding soundness evaluation system. *Vet. Clin. N. Am. Food Anim. Pract.* 13(2):283-93.

Kobayashi, H., Urashima, M., Hoshi, Y., Uchiyama, H., Fujisawa, K., Akatsuka, J., Maekawa, K. and Hurusato, M. 1996. Testicular morphological changes in children with acute lymphoblastic leukemia following chemotherapy. *Acta Paediatr. Jpn* 38(6):640-3.

Koski, W.R., Davis, R.A., Miller, G.W. and Withrow, D.E. 1993. Reproduction. pp. 239-74. *In:* J.J. Burns, J.J. Montague and C.J. Cowles (eds.) *The Bowhead Whale*. 1st. Edn. Society for Marine Mammalogy, Lawrence, Kansas. Special Publication. No.2. 787pp.

Lephart, E.D. and Watson, M.A. 1999. Maternal separation: hypothalmic-preoptic area and hippocampal calbindin-D28K and calretinin in male and female infant rats. *Neurosci. Lett.* 267(1):41-4.

Lephart, E.D., Taylor, H., Jacobson, N.A. and Watson, M.A. 1998. Calretinin and calbindin-D28K in male rats during postnatal development. *Neurobiol. Aging* 19(3):253-7.

Meachem, S.J., Neischlag, E. and Simoni, M. 2001. Inhibin B in male reproduction: pathophysiology and clinical relevance. *Eur. J. Endocrinol.* 145(5):561-71.

Mogoe, T., Takatoshi, S., Asada, M. and Fukui, Y. 2000. Functional reduction of the southern minke whale (*Balaenoptera acutorostrata*) testis during the feeding season. *Mar. Mammal Sci.* 16(3):559-69.

Nadji, M. and Morales, A.R. 1983. Immunoperoxidase: Part I. The technique and its pitfalls. *Lab. Med.* 14(12):767-71.

O'Hara, T.M., Albert, T.F., Oen, E.O., Philo, L.M., George, J.C. and Ingling, A.L. 1999. The role of Eskimo hunters, veterinarians, and other biologists in improving the humane aspects of the subsistence harvest of bowhead whales. *J. Am. Vet. Med. Assoc.* 214(8):1193-8.

O'Hara, T.M., George, J.C., Tarpley, R.J., Burek, K. and Suydam, R.S. 2002. Sexual maturation in male bowhead whales (*Balaena mysticetus*) of the Bering-Chukchi-Beaufort Seas stock. *J. Cetacean Res. Manage.* 4(2):143-48.

Short, R.V. 1997. The testis, the witness of the mating system; the site of mutation and the engine of desire. *Acta Paediatr. Suppl.* 422:3-7.

Tarpley, R.J., Jarrell, G.H., George, J.C., Cubbage, J. and Stott, G.G. 1995. Male pseudohermaphroditism in the bowhead whale, *Balaena mysticetus*. *J. Mammal.* 76(4):1267-75.

Vleigen, M.K., Schlatt, S., Weinbauer, G.F., Bergmann, M., Groome, N.P. and Nieschlag, E. 1993. Localization of inhibin/activin subunits in the testis of adult nonhuman primates and men. *Cell Tissue Res.* 273:261-8.

Evaluation of nutritive condition and reproductive status of migrating gray whales (*Eschrichtius robustus*) based on analysis of photogrammetric data

WAYNE L. PERRYMAN AND MORGAN S. LYNN

Southwest Fisheries Science Center, NMFS, NOAA, PO Box 271, La Jolla, California 92038, USA
Contact e-mail: wayne.perryman@noaa.gov

ABSTRACT

Vertical aerial photographs were collected of gray whales migrating along the California Coast between 1994 and 1998 to readdress some published findings on the biology and life history of this population based on examination of specimens. For each whale, an attempt was made to measure standard total length, the width of the whale at its widest point, the distance from the tip of the rostrum to the widest point, and the width of the flukes. For southbound gray whales, early migrants were longer on average and more likely to be parturient than those migrating later. Near-term pregnant females were wider relative to their length than other southbound gray whales. This difference was easily detected by visual inspection of the images and through statistical evaluation of length and width data. There was 100% agreement between identification of parturient females based on linear regression analysis of length and width and discriminate analysis of all measurements. Based on the proportion of parturient females to those with calves during sampling of southbound whales, the median calving date was estimated to be 13 January. Southbound calves averaged 4.6m in length; those photographed northbound in late April, at an age of about three months, averaged 7.1m. Average length for yearlings, based on combined southbound and northbound data, was 8.5m. Residuals from a regression of width on length were compared, and significant changes in the relationship were detected which were consistent with changes in nutritive condition or fatness described from examination of whales taken along the California Coast between 1959 and 1969 (Rice and Wolman, 1971). Parturient females were the widest relative to their length and northbound cows with calves were the narrowest in the sample. The relationship between length and width for migrating gray whales that were not parturient or associated with a calf, showed that southbound gray whales were significantly wider than northbound whales photographed approximately 60 days later. These results indicate that the predictable but relatively small changes in condition or fatness of gray whales associated with fasting during their winter migration can be reliably detected in measurements from vertical aerial photographs.

KEYWORDS: GRAY WHALE; PHOTOGRAMMETRY; MORPHOMETRICS; PREGNANCY; GROWTH; CONDITION

INTRODUCTION

As with most species of large cetaceans, our understanding of the basic biology and life history of the eastern Pacific population of gray whales (*Eschrichtius robustus*) is derived mainly from the examination of animals killed in fisheries. For this population, most specimens have come from the Chukotkan native fishery, which has taken gray whales in the Bering and Chukchi Seas since the mid-1930s and those animals taken under special research permits along the California coast between 1959 and 1969. Examinations of whales from the Chukotkan fishery have provided much of the basic knowledge on the growth, physiology, reproduction and food habits for this population (Zimushko and Ivashin, 1980; Bogoslovskaya *et al.*, 1981; Blokhin, 1984; Yablokov and Bogoslovskaya, 1984; Blokhin and Tiupeleev, 1987). Recently, most of the whales from this fishery have been taken from nearshore waters, and the non-random distribution of whales on the feeding grounds has biased this sample towards immature specimens and females.

In addition, 316 gray whales were taken under seven scientific research permits issued to the US Bureau of Commercial Fisheries (now the National Marine Fisheries Service) under Article 8 of the Convention of the International Whaling Commission between 1959 and 1969. Land-based catcher boats were used to capture whales which were brought to the Richmond Shore Stations (inside San Francisco Bay) in California, where external measurements and biological samples were collected. These takes were conducted to provide representative samples from both the southbound and northbound migrations. The monograph by Rice and Wolman (1971) presenting the results of this study remains the most comprehensive review of the biology and ecology of this population. Among the many findings on basic anatomy, growth and reproduction that complement the Russian studies, these authors reported partial temporal segregation by age, sex and reproductive status in migrating gray whales. They also found that gray whales experience a weight loss of between 11-25% during approximately 60 days of fasting between their southbound and northbound migrations past central California. They reported that changes in girth were better indicators of this change in condition than measurements of blubber thickness. Although they planned to sample all segments of the population, no northbound cows with calves were taken. In addition, the authors recognised that their analyses were based on data from one source, with some known biases (i.e. selection of larger whales by gunners) and they recommended that their findings be re-examined through other approaches with different biases (Rice and Wolman, 1971; Rice, 1990).

The data available from the Russian and scientific catch are supplemented by information from stranded gray whales, a few that have been captured live, as well as a small set of body measurements made from photographs. Although most of these samples support the findings from the studies involving direct takes, there are some notable exceptions. For instance, size at birth estimated from stranded calves, with the exception of a very recent sample (Pacheco, 1998), is generally smaller than that taken from near-term foetuses or calves captured live (Norris and Gentry, 1974; Bryant *et al.*, 1984; Jones and Swartz, 1984; Sumich, 1986). In addition, growth rates for young gray whales based on stranded and captive whales indicate that growth is much slower than that projected from fishery data (Sumich, 1986). The reason for this apparent difference in estimated growth

remains unresolved, although it may be related to the possibility that weaker, slower growing calves are more prone to stranding.

This paper presents the results of an aerial photogrammetric sampling programme that was modelled after the effort conducted more than 25 years ago by Rice and Wolman (1971). Migrating gray whales were photographed as they passed along the California coast and the field effort was scheduled in an attempt to collect representative samples of the southbound and northbound migrations. The field season was extended into April and May to include the northbound gray whale cows with calves. Measurements from vertical photographs of gray whales were examined for evidence of temporal segregation during the migrations, for indications of reproductive status, and for changes in shape that may reflect reduction in nutritive condition during the winter migration. The goals of the study were to determine whether data from high resolution aerial photographs could be used to detect changes in nutritive and reproductive condition in gray whales, and to review other findings based on examination of fishery and stranded specimens.

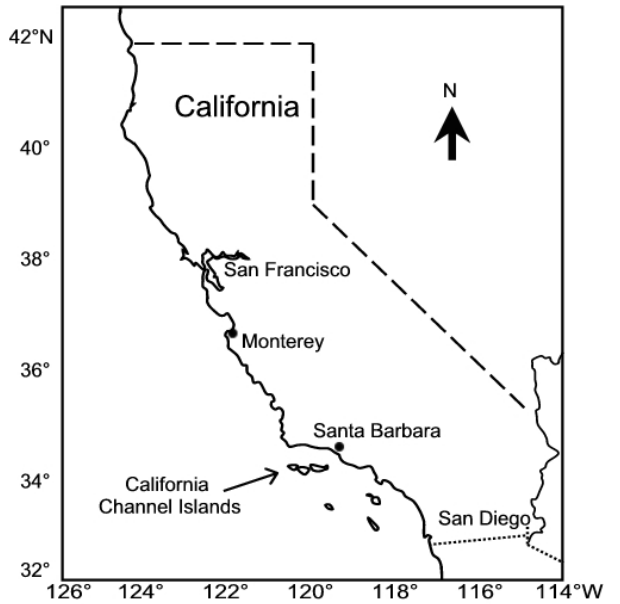


Fig. 1. Aerial photographs of gray whales were taken around the California Channel Islands and along the central California coast between Santa Barbara and Monterey.

MATERIALS AND METHODS

The measurements presented in this paper were made on vertical aerial photographs of gray whales taken near the coast between Monterey, California and the California Channel Islands (Fig. 1). Photographing of gray whale cows and calves began during the northbound migration in 1994 and continued each spring through 1998. In 1996, sampling of northbound gray whales was expanded to include the adults and juveniles that comprise the first phase of this migration; this photographic sampling continued in 1997 and 1998. In the winters of 1997 and 1998, gray whales were also photographed during the southbound migration. Since migrating gray whales reportedly segregate by age, sex and reproductive condition during the southbound and the first phase of the northbound migrations (Rice and Wolman, 1971; Poole, 1984), the field effort was scheduled to span the peaks of these migrations (Table 1). Although the southbound migration is characterised by a steep and predictable peak in migration rates (Rugh *et al.*, 2001), the northbound migration is much more diffused and selection of a peak date can be somewhat arbitrary (Pike, 1962; Herzing and Mate, 1984; Poole, 1984). The northbound peak date for this study was selected as 14 March based on the studies cited above and the observations of gray whale migration rates within the Southern California Bight.

Aerial photographs were taken with a 126mm format KA76 military reconnaissance camera (image size 114mm × 114mm) that was mounted vertically over a hatch in the deck of a *Partenavia* aircraft. A bubble level was attached to the top of the camera and the orientation of the camera was adjusted during each pass to ensure that the photographs were collected vertically. When the camera fired, a data acquisition system recorded an altitude reading from a *Sperry* 300 series twin transducer radar altimeter, and a position from a GPS unit. Gray whales were photographed from altitudes between 135 and 245m. The majority of the photographs (about 98%) were taken with *Kodak* Aerecon Plus-X aerial film. Two colour transparency films, *Kodak* MS 2448 and *Kodak* HS SO359, were tested early in the experiment and the resolution of the black and white negatives was consistently superior to the colour transparencies.

Table 1

Dates of aerial photographic missions to sample southbound and the first phase of the northbound migrations of gray whales. This sampling was conducted around the California Channel Islands.

Migration	Sampling periods		
	Early	Peak	Late
Southbound migration	6, 7 Jan. 1997 30,31 Dec. 1997	17, 18 Jan. 1997 17 Jan. 1998	29, 30 Jan. 1997 23, 24 Jan. 1998
Northbound migration (Phase 1)	11 Mar. 1998	14, 15 Mar. 1996	25, 26, 27 Mar. 1996 18, 19 Mar. 1997 20 Mar. 1998

Measurements from photographs

The black and white negatives were reviewed on a light table and the whales selected for measurement were those clearly visible and swimming near the surface. If acceptable photographs were available from more than one pass over a pod of whales, natural marks were used to identify individuals and thus avoid inclusion of duplicate measurements from the same whale. A high resolution digital camera (*Diagnostic Instruments* SPOT) mounted above the light table was used to capture images of individual whales which were then transferred to a computer. The whales were measured with either *Image*, a software package developed by the *National Institutes of Health*, or *Image Pro Plus*, produced by *Media Cybernetics*. The distances measured on the computer screen were multiplied by the scale of the photograph (scale = altitude/lens focal length) to convert them to true or ground distances. Because data from radar altimeters are generally precise but may be biased, calibration experiments were conducted each season to determine the sign and magnitude of any bias in the recorded altitude readings. In these experiments, several photographic passes were made over pipes of known length that were towed offshore. These experiments were conducted over water to avoid the small positive bias that has been reported from calibration experiments conducted over targets on land (Koski *et al.*, 1992). Correction factors were then developed for recorded

altitudes based on linear regression of recorded altitude and altitude calculated from target measurements, as shown in Fig. 2 (Perryman and Lynn, 1993; Gilpatrick, 1996; Perryman and Westlake, 1998).

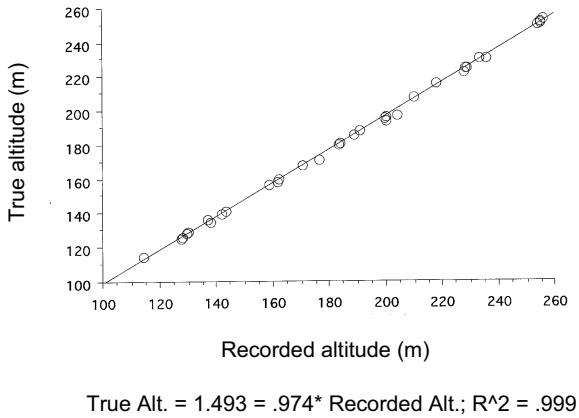


Fig. 2. The altimeter calibration equation calculated for altimeter readings recorded during photographic missions flown in 1997. True altitudes were determined by measurements of targets of known length.

For each whale, an attempt was made to measure total length (L_t); the width of the whale at the widest point (W_m); the distance from the tip of the rostrum to the widest point (RW_m); and fluke width (F_w) as shown in Fig. 3. After making several photographic passes over a whale, many excellent images were often achieved in which these characters could be accurately measured. The following criteria were used for selecting the best measurements. Because the normal movements associated with swimming can make L_t, RW_m, and F_w appear smaller when seen from above, the largest measurement was selected when more than one measurement was obtained from the photographs. Measurements of W_m can be either positively or negatively biased as the whale bends its body to begin a dive or from distortion caused by waves created as a whale slides along the surface. For this feature, the measurement was selected from the photograph in which the body of the whale was not flexed and both sides of the whale were most distinct.

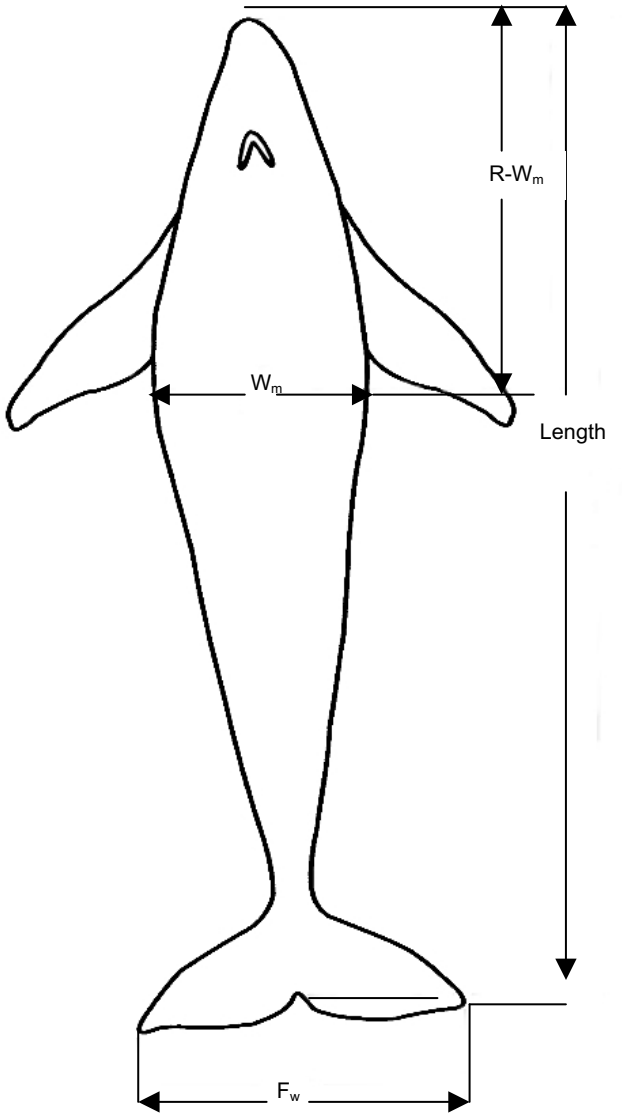


Fig. 3. Features measured on vertical photographs of migrating gray whales.

Accuracy and precision

A common concern in any study based on measurements of continuous variables is whether the measurements are accurate (reflective of the true distance measured) and precise (differences between repeated measures of the same feature on the same individual are small). To test the accuracy of the measurements, two 7.6cm diameter sections of plastic pipe were towed offshore and photographed from the range of altitudes used while photographing gray whales. The lengths of the two pipes were approximately that of a cow (12.02m) and a calf (6.02m). The pipes were measured with the same image capture and measurement system used for measuring whales. All measurements of these targets were performed by individuals who did not know the true lengths of the targets. The average lengths from the target photographs were 12.04m ($n = 41$; $s = 0.169$) and 6.04m ($n = 41$; $s = 0.093$). Measurement accuracy in the test was within 1% for each pipe and coefficients of variation (CV) for the two sets of measurements were 0.014 and 0.015 for the longer and shorter pipes respectively.

Although the target measurements indicate that lengths of semi-rigid objects floating at or near the surface can be determined within an acceptable degree of accuracy and precision, these data do not replicate the additional variance found in repeated measures of mammals that flex their bodies as they swim. To test the precision of the measurements of whales, the coefficients of variation (CV) were calculated for a series of whales that had been photographed several times. CVs were only calculated for whales that were measured on at least three photographs. For L_t, W_m and F_w, the averages of the CVs were 0.020, 0.035 and 0.030 respectively. The level of sampling precision and tendency for measurement error to represent a larger proportion of shorter measurements are consistent with other photogrammetric studies on large cetaceans (Best and Rüther, 1992; Koski *et al.*, 1992; Angliss *et al.*, 1995).

Analyses

The measurements of adult and juvenile gray whales were divided into strata based on whether they were photographed before, during, or after the peak of the migrations (Table 1). The means of the lengths of whales from these strata were then compared to look for evidence of the partial temporal segregation in migrating whales reported by Rice and Wolman (1971). For instance, asymptotic length for female gray whales is larger than for males, and the average length

for a sample of whales that included a disproportionate number of adult females would likely be longer than the average of a sample composed of mostly adult males or a combination of males and older juveniles. In tests of the hypothesis of homogeneity in length between strata, observations were truncated at 10m to eliminate the few very small juveniles that were found in some strata. Tukey/Kramer tests were used for post hoc comparisons, because the power of this test remains high when sample sizes are unequal (Day and Quinn, 1989).

The following assumptions were made about the age, sex and reproductive condition of individual whales based on their size, association with a smaller whale or their shape. It was assumed that all whales < 8m in length and swimming in close association with a large whale (> 11m) were calves. The larger whale swimming in close association with a calf was identified as the mother of that calf, and is referred to as a cow or an adult female. Some whales were also classified as parturient or near-term pregnant females based on the relationship between their length and width. These animals are referred to as 'parturient females' in the text to distinguish them from other females that may have recently become pregnant.

It was noted that the ratio of parturient to post-partum female gray whales changed as the southbound migration progressed. A logistic model was fitted to the proportion of those female gray whales identified as reproductive (cows with calves and parturient females) that were post-partum for missions flown on 30-31 December, 6-7 January, 17-18 January and after January 22nd. From this, the median birth date for this population was estimated. The confidence intervals for this estimate were estimated by anchoring the model at 100% for the last point, bootstrapping each of the other points and refitting the model 999 times. This estimate of median birth date is dependent upon the assumption that calving is dependent upon the duration of the pregnancy, not the geographic location of the near-term pregnant female. This assumption is supported by recent findings suggesting a relationship between the shift towards later migrations in recent years and an increase in sightings of calves along the California coast (Shelden *et al.*, 1997; Rugh *et al.*, 2001; Buckland and Breiwick, 2002).

The measurements included data that reflect whale size (L_t) and shape (W_m, RW_m). In order to compare sets of shape measures between groups of whales that differ in size, it was first necessary to remove the effect of size on shape. A least squares regression technique was selected that has been shown to be very effective in removing the effect of size in morphometric studies (Atchley *et al.*, 1976; Reist, 1985) and in evaluation of condition indices for harbour porpoise (Read, 1990). A linear regression of each shape parameter was performed on length and the resulting residuals from these regressions were used as size-free variates for normal hypothesis testing. Tests were conducted between groups of whales that were identified based on their direction of migration or reproductive condition (cows and parturient females).

The main reason for comparing measures of shape was to determine whether documented seasonal changes in gray whale nutritive condition could be detected in measurements from photographs. Gray whales rely on the oxidation of stored fats to support their metabolic needs during the migration, and thus are thinner when they return to their feeding grounds in the Arctic than when they departed (Rice and Wolman, 1971; Blokhin, 1984). Rice and Wolman (1971) reported that the changes in condition from fatter southbound to thinner northbound gray whales were best reflected in changes in girth. W_m was used as the proxy for girth and an investigation of whether changes in condition were reflected in changes in the relationship between length and width was calculated. The location of the point of maximum width (RW_m) was also compared to determine whether this measure varies with reproductive or nutritive condition.

Several multivariate techniques were used during the exploratory phase of the analyses, but only the results of one discriminant analysis are reported in this paper. Multivariate analyses were performed in *MiniTab* (release 10) and all other analyses were done in *Statview* (version 5). Differences were considered statistically significant if $P < 0.05$.

RESULTS

Southbound gray whales

The lengths of 303 southbound gray whales were measured from the photographs collected in 1997 and 1998. The measurements were divided into three strata based on the timing of the photographic missions relative to the peak of the southbound migration (Fig. 4). Calves of the year were found in all three strata and represented about 5% of the total length sample. These data probably overestimate the proportion of calves in the southbound migration because cows with calves spend more time at the surface and were thus more easily detected from the air. Average length for the southbound calves was 4.6m ($n = 15$; range 4.3-4.9m), and average length for the accompanying cows was 12.2m ($n = 15$; range 11.4-13.0m). No significant correlation was found between the length of a cow and her calf ($r = 0.261$; $P = 0.35$). Unlike calves, the very small juvenile gray whales which averaged 8.5m ($n = 5$; range 8.1-8.8m) were all photographed late in the migration. The length samples in the three strata were truncated at 10m; the null hypothesis that the means of the three strata did not differ was tested and rejected ($F = 9.948$, $P < 0.001$). *Post hoc* tests found that whales photographed prior to the peak of the migration were longer on average than those from the peak or post-peak strata (Table 2).

While reviewing the photographs of southbound gray whales, it was noted that some of the longer whales (> 11m) were exceptionally wide relative to their length (Fig. 5). These wide-bodied whales were most common early in the southbound migration (Fig. 6) and were suspected by the authors to be near-term pregnant females. A technique commonly used to identify outliers in regression analysis was selected to separate these parturient females from the remainder of the sample.

L_t and W_m measurements for whales photographed late in the migration (when no exceptionally wide whales were detected) and those < 11m in length were used to build a linear model for the length and width relationship for whales that fall within the main cluster of points (i.e. not exceptionally wide) shown in Fig. 6. Measurements of whales swimming with a calf were excluded from this exercise. The standard deviation of the residuals for this regression of L_t (x) and W_m (y) was calculated and then the relationship between width and length was compared for the southbound whales not used to construct the model. It was assumed that any whale whose L_t and W_m measurements produced a residual that was greater than 1.99 ($t_{(0.05)(85)}$) standard deviations from the above regression was an outlier (parturient). A total of 34 southbound gray whales were classified as parturient females based on their large positive

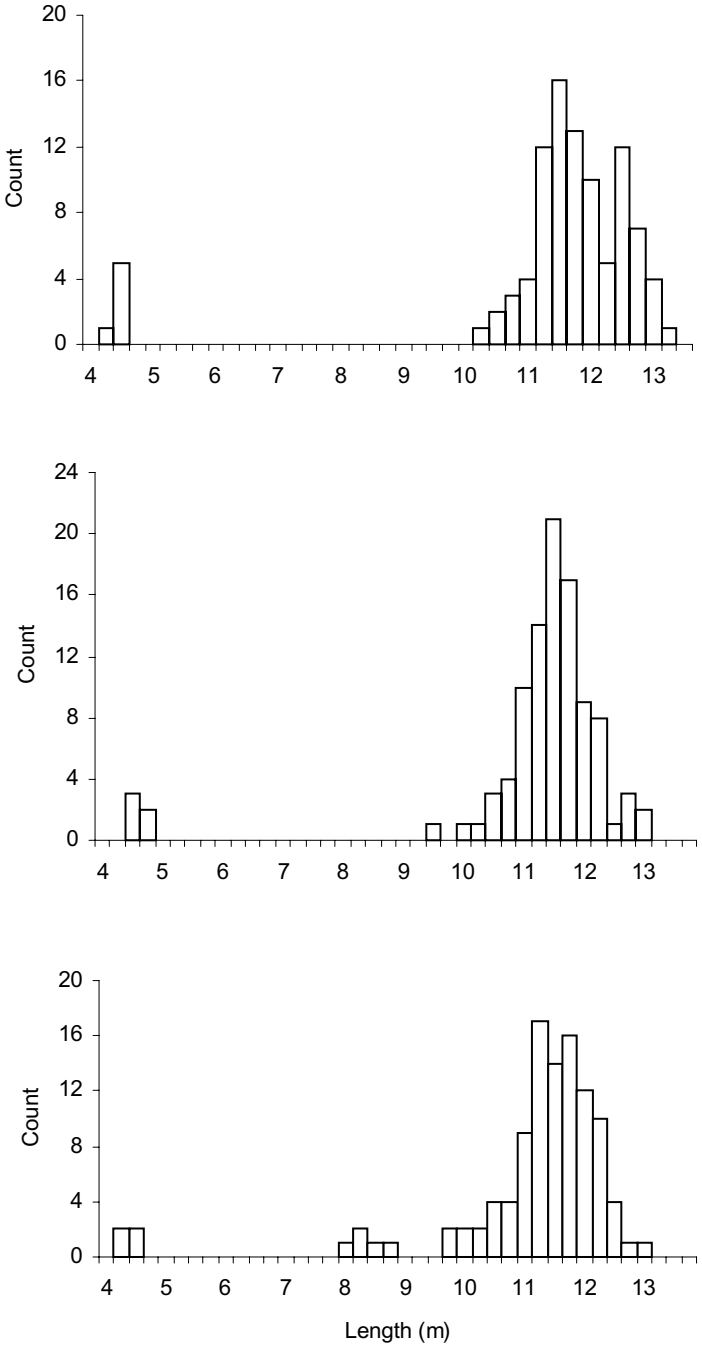


Fig. 4. Histograms of lengths for whales photographed early (12/30-1/7), near the peak (1/17-1/18), and late (1/23-1/30) in southbound migrations of 1997 and 1998.

Table 2

Results of Tukey/Kramer *post hoc* comparisons between means of lengths for gray whales >10m photographed early, near the peak and late in the southbound migration.

	Sample size	Mean (m)	Comparisons	Differences	Critical difference	P<0.05
Early	90	11.95	Early *vs* peak	0.296	0.260	Yes
Peak	95	11.65	Early *vs* late	0.479	0.255	Yes
Late	103	11.47	Peak *vs* late	0.183	0.251	No

residuals. For both the parturient and the remaining southbound gray whales sampled, there was a positive linear relationship between length and width (parturient females $F = 10.95$, $P < 0.01$; remaining southbound $F = 115.09$, $P < 0.01$) (Fig. 7).

Inconsistencies in the classification of parturient females based on L_t and W_m were investigated by performing a discriminant analysis on the full set of measurements (L_t, W_m, RW_m and F_w) for all southbound whales >10m. A cross-validation technique was used in the analysis to reduce the probability of overly optimistic matches with the original classifications. There was 100% agreement between classifications based on the regression technique and discriminant analysis.

In late December, about 47% of the southbound gray whales photographed were either 'parturient' or accompanied by a calf. By the end of January, parturient females and females with a calf decreased to 2% of the photographic sample. While the number of parturient females and cows decreased, the proportion post-partum (cows with calves/parturient females + cows with calves) increased steadily as the migration progressed. The data were divided into four samples (30-31 December, 6-7 January, 17-18 January and after 22 January) and the proportion post-partum was calculated for each. A logistic model was fit to these four points (Fig. 8) and it was estimated that the median birth date for parturient gray whales passing through the Channel Islands was 13 January (95% confidence intervals 12-15 January).

Northbound gray whales

Phase 1- adults and juveniles

Nine missions were flown to sample the adult/juvenile phase of the northbound migration; 273 whales were measured from the photographs. The data were divided into three strata based on the dates of the photographic mission relative to an estimated migration peak of 14 March (Fig. 9). Length strata were truncated, at 10m and the null hypothesis that the means of the samples were equal was tested. Tests revealed no difference in average length for whales photographed before, during, or after the peak of the northbound migration ($F = 0.108$, $P = 0.898$). The relationship between length and width was consistent across strata. The six small juveniles, mean length 8.5m (range 8.29-8.93m) photographed (including one from Phase 2) were found near the peak or late in the migration. No cow-calf pairs were found in this phase of the migration.

Phase 2 —cows and calves

Northbound gray whale cows and calves were photographed between mid-April and early May each year from 1994-1998 (Figs 10a and 10b). Calves averaged 7.10m in length ($n = 112$) and ranged from 5.5-8.1m. Cows averaged 12.4m ($n = 98$) and ranged from 11.2-13.9m. A significant positive linear relationship ($b = 0.245$, $p = 0.002$) was found between the length of a cow and the associated calf (Fig. 11). Thus, longer cows were associated with longer calves.

Adult females

Due to their shape (southbound whales only) or association with a calf, 49 southbound and 98 northbound gray whales were classified as adult females. Southbound adult females were somewhat shorter on average (12.1m) than those photographed northbound (12.4m) (t-test, $P = 0.012$).

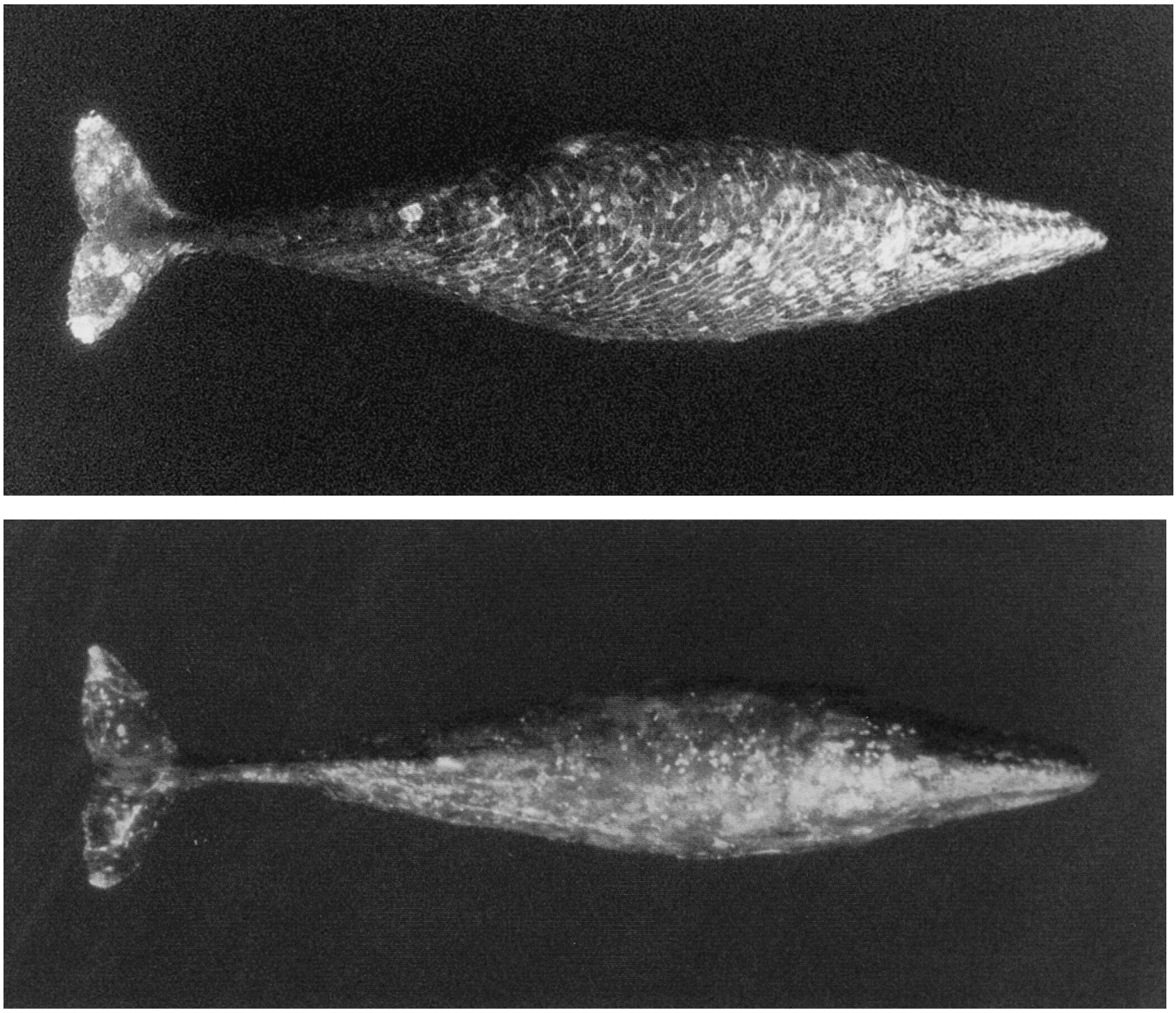


Fig. 5. Vertical photographs of two southbound gray whales illustrating the difference in shape between very wide whales and those more fusiform in shape.

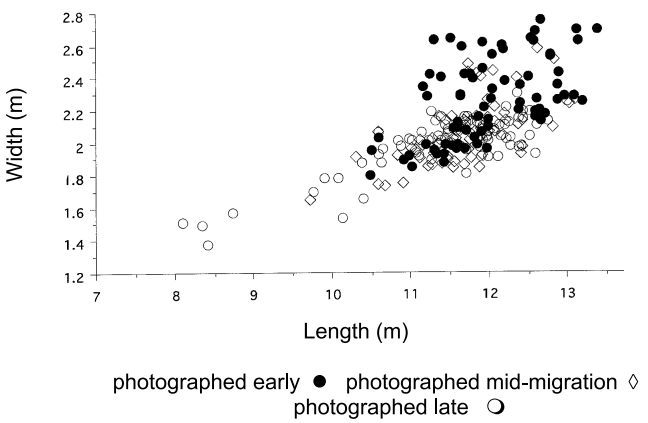


Fig. 6. Plot of length and width of southbound whales photographed early (●), near the peak (◇) and late (○) in the southbound migrations of 1997 and 1998.

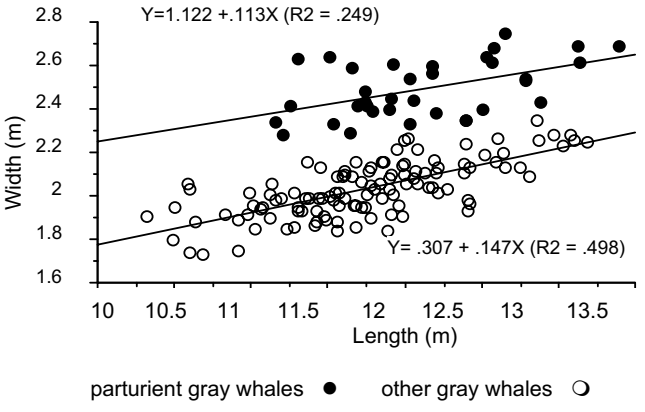


Fig. 7. Plot of length and width measurements for parturient (●) and other southbound gray whales (○). Measurements of southbound gray whales accompanied by calves were excluded from this dataset. Lines represent least squares linear regression fits to each dataset.

Possibly younger, shorter females that were parturient or had given birth during the southbound migration are more likely to lose their calf and thus were not classified as cows in the northbound sample.

Length at one year

Length distributions from measurements of south and northbound gray whales (Fig. 12) included a group of very small juvenile whales that were separate from the main

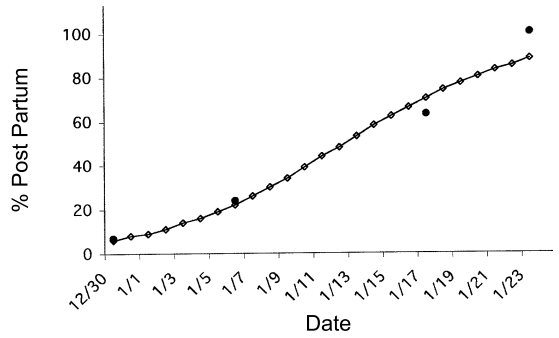
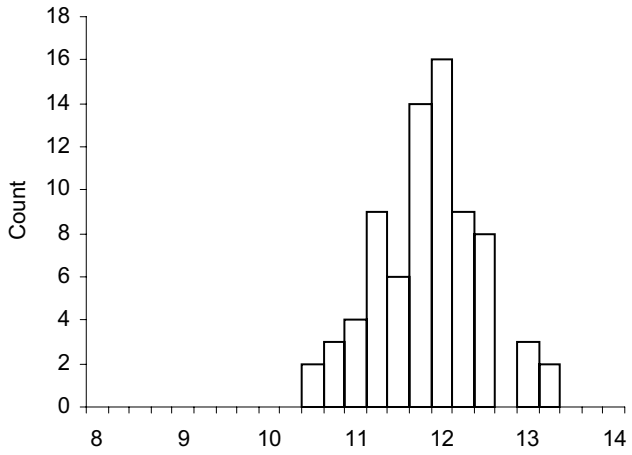
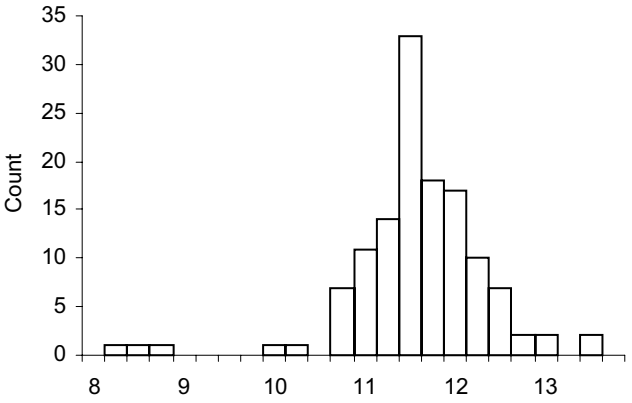
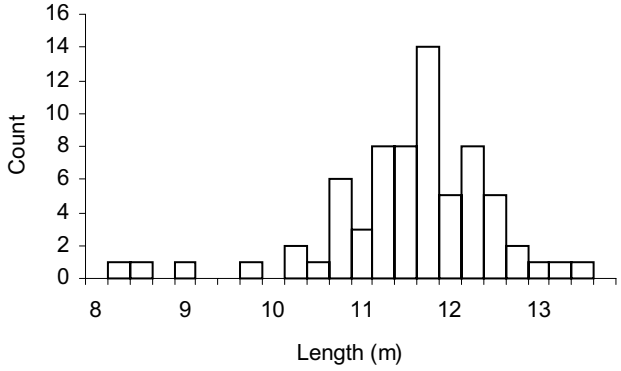


Fig. 8. Fit of a logistic model to the fraction of reproductive female gray whales (parturient and cows with calves) that were post-partum for the four sampling periods in 1997 and 1998.

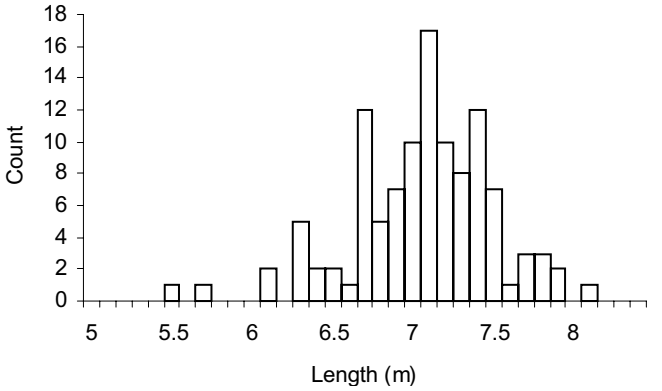






early = 3/11 peak = 3/14-3/18 late = 3/20-3/25

Fig. 9. Histograms of lengths for whales photographed early (3/11), near the peak (3/14-3/18), or late (3/20-3/25) in northbound migrations of 1996-1998.



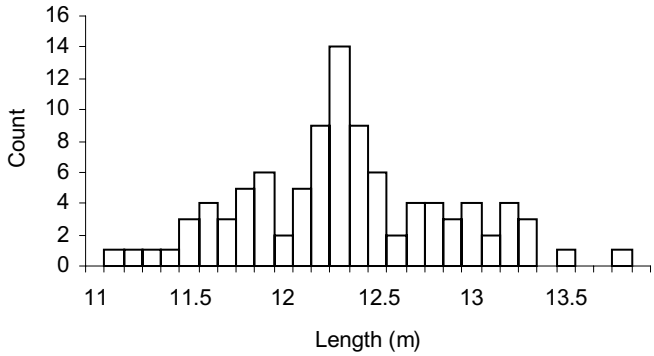


Fig. 10a and b. Histograms of lengths of calves (a) and cows (b) photographed migrating northbound along the central California coast between 1994 and 1998.

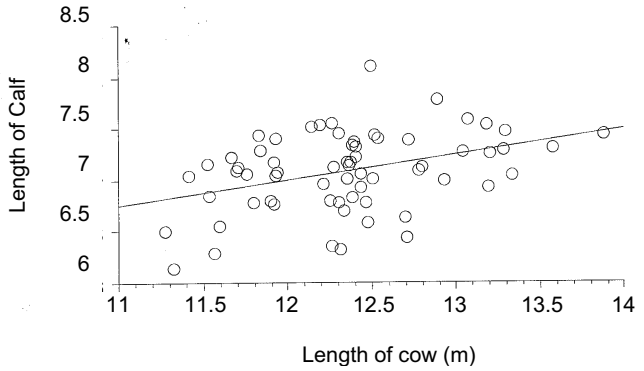


Fig. 11. Relationship between length of northbound cows and their associated calves measured from vertical aerial photographs. Line in the figure is the least squares linear regression fit to these data.

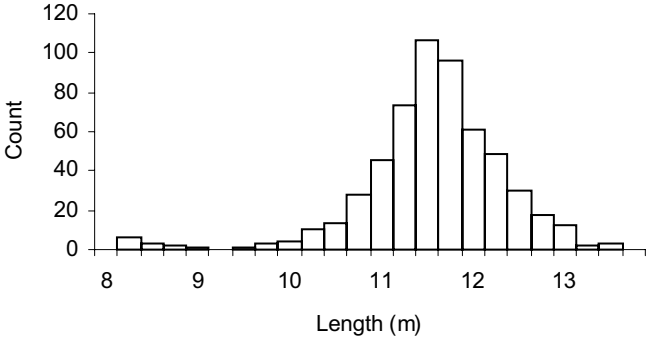


Fig. 12. Lengths of all southbound and northbound gray whales, with the exception of calves of the year, measured from vertical aerial photographs.

distributions. These small juveniles represented about 2% of the total dataset, and were probably yearling gray whales. The average length for each sample, and thus the combined sample, was 8.5m (n = 11; range 8.1-8.9). The categorisation of these yearlings was based on the assumption that the hiatus in length distribution represents the difference in length between one- and two-year-old gray whales. The sample probably underestimates the proportion of one-year-olds in the migration because these whales were small, difficult to detect from the air, and were generally found swimming alone.

Nutritive condition-fatness

For all whales (except calves), a linear regression was fitted to the length and width measurements and the residuals of this least squares fit were saved as derived shape variables. The means of the residuals (ANOVA) for parturient females, southbound cows, southbound adults and juveniles, northbound adults and juveniles, and northbound cows were compared and the null hypothesis that they were the same was rejected (F = 317.8, P < 0.001). Tukey/Kramer *post hoc* tests revealed that all samples differed significantly (Fig. 13). These results indicate that gray whale condition, as indexed by the relationship between L_t and W_m, declined steadily during the winter migration and that northbound lactating females were the narrowest or in the poorest nutritive condition of all categories.

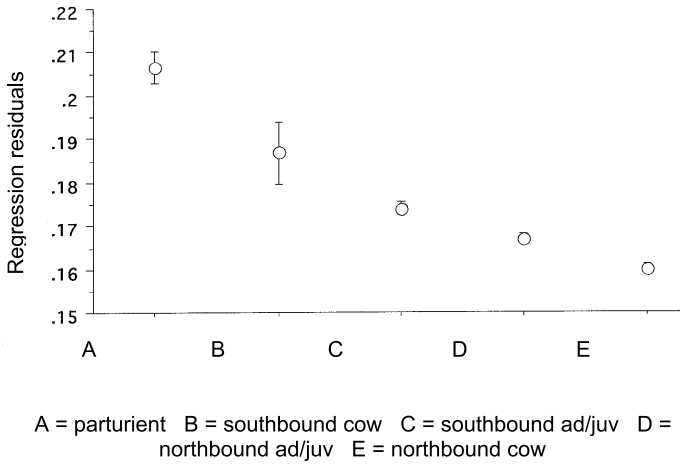


Fig. 13. Plot of means of residuals from regressions of length (x) on width (y) for all gray whales for which these two features could be measured. Calves of the year were excluded from this dataset. Error bars are + 2 standard errors of the mean values.

To test the hypothesis that RW_m was the same between the five categories of whales described in the previous paragraph, a linear regression of RW_m on L_t was performed. The means of the resulting residuals were then compared (ANOVA). The location of W_m was found to vary significantly between whale categories (F = 40.046, P < 0.001). Results of Tukey/Kramer *post hoc* tests indicated that for parturient females and southbound cows (the two categories of whales that were widest relative to their length) the location of their greatest width was farther from their rostrum than for all other categories of gray whales (Fig. 14). For the near-term pregnant females the foetus probably accounts for some of this change in shape, but for southbound cows these data indicate that the extra fat carried by these females was reflected in a change in width that is closer to the fluke than in other southbound or northbound whales.

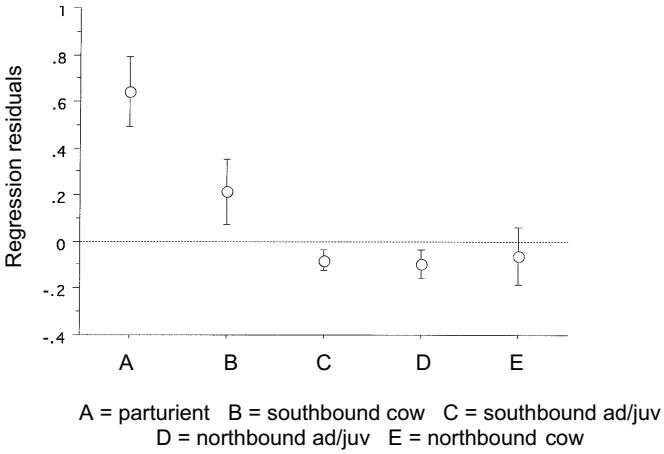


Fig. 14. Plots of means of residuals from regression of length (x) on distance from the tip of the rostrum to the widest point (y) on all gray whales for which these two features could be measured. Calves of the year were excluded from this dataset. Error bars are + 2 standard errors of the mean values.

DISCUSSION

The findings of partial temporal segregation in southbound gray whales by size and reproductive condition are consistent with those published by Rice and Wolman (1971) from analysis of whales sampled through a hunt over 25 years earlier. Early migrants were longer on average and more likely to be parturient than those passing later. Small juveniles were most common after the peak of the migration. Although no southbound cow/calf pairs were located during Rice and Wolman's (1971) study, in this study calves represented about 5% of the southbound sample. Shelden *et al.* (1997) found that up to 4.4% of the southbound gray whales sighted during aerial surveys near Monterey Bay were neonates and suggested that the apparent increase in births north of the lagoons may be associated with the delays in the migration reported over the past few decades (Rugh *et al.*, 2001; Buckland and Breiwick, 2002).

The median calving date within the Southern California Bight was estimated as 13 January. This date falls between the median calving date estimates of 10 January by Rice and Wolman (1971) based on foetal growth rates, and 27 January proposed by Rice *et al.* (1981) from the temporal distribution of calves in Laguna Ojo de Liebre, Mexico. Although the estimate 13 January date for gray whales is the only estimate based on counts of parturient and recent post partum females, its validity for the eastern North Pacific population depends on the untested assumption that the probability of a near-term female giving birth is more related to the duration of her pregnancy than to her location.

Making the assumption that a female mammal is pregnant, based on her shape alone, can be risky. However, the identification of parturient female gray whales from the relationship between their length and maximum width is supported by several findings made from the examination of specimens taken in fisheries. Fishery data revealed that near-term pregnant gray whales had greater girth to length ratios than any other group of southbound gray whales and that pregnant females were most common early in the

southbound migration (Rice and Wolman, 1971). In addition, results from examination of gray whales taken in the Arctic and off California agree that female gray whales reach sexual maturity at a length of about 11.1m (Rice and Wolman, 1971; Blokhin, 1984; Yablokov and Bogoslovskaya, 1984). This study found that some of the southbound gray whales photographed were exceptionally wide relative to their length, these wide whales were most common early in the southbound migration, and that all of these whales were >11.1m in length. In addition, no exceptionally wide whales were found in any of the photographs of gray whales returning northward from the calving lagoons. Because of the consistency of the findings from the fishery and the photogrammetric sample, it is likely that the anomalously wide whales photographed were near-term pregnant females.

While no pattern of temporal segregation by size and sex was found in the northbound migration, as reported from specimens, the early migration sample may have been collected too late in the season to capture the northbound newly pregnant females. Rice and Wolman (1971) noted that these females pass Central California within a two-week period that peaks around 28 February and, although gray whales appear to be migrating later than they did in the 1960s (Rugh *et al.*, 2001; Buckland and Breiwick, 2002), the March 11 flight may have missed this first pulse of northbound whales. No cows with calves were found to migrate with the adults and juveniles that make up the first phase of the northbound migration, as also reported by Poole (1984).

Although no correlation was found between the lengths of southbound gray whale cows and their associated calves, there was a positive linear relationship between the lengths of northbound cows and calves. These results suggest that length at birth is independent of size of the mother, but calves of larger (probably older) cows grow at a faster rate than those of smaller females. However, southbound data include measurements of only 15 cow/calf pairs and the lack of correlation in these pairs of measurements may be the result of the small sample size. The average length for the 15 newborn calves (4.6m) is closer to estimates of size at birth based on measurement of near-term foetuses (4.6m) from Rice and Wolman (1971) and of live-captured calves (4.7m) (Norris and Gentry, 1974) than it is to the average of lengths for stranded calves (4.4m) summarised by Jones and Swartz (1984). Rice (1983) suggested that compression of the intervertebral disks might cause body length to shrink after birth, thus explaining the difference between data from foetuses and stranded newborns. The results here and recent data from stranded calves (Pacheco, 1998) indicate that the difference between these datasets is more likely the result of higher mortalities (and therefore strandings) amongst smaller, possibly premature calves. A similar pattern of potential size bias towards small calves in data from strandings was reported for southern right whales (Best and Rüther, 1992).

Based on their analysis of the data taken from the whales captured along the California coast, Rice and Wolman (1971) estimated that gray whales reached a length of about 8.5m at weaning and averaged 9.3m by the following winter. Their estimate of length for one-year-old whales was based on the assumption that the smallest whales in the sample were yearlings. These small gray whales had two growth layer groups in their ear plugs, and more recent research has shown that gray whales lay down a single layer each year (Blokhin and Tiupeleev, 1987). Sumich (1986) reviewed data from stranded specimens, some photogrammetric data and growth of a captive gray whale calf and concluded that gray whales reach a length of 5.6m at three months, are 7m at weaning and attain a length of 8.0m at one year. Sumich (1986) assumed that the hiatus in lengths of stranded neonates in the 6-7m range represented the difference in lengths between calves of the year and one-year-old whales. In a more recent study, Sumich *et al.* (2001) published growth data for a second gray whale calf that was raised in captivity. This calf reached a length of about 6m at three months and 9m at one year.

In this study, gray whale calves photographed in late April, at an age of about three months, averaged 7.1m in length, not 5.6-6m as suggested above. The lack of stranded specimens in the 6-7m size range is likely to reflect greater survivorship of calves that reach this size. Based on data from 9-10 month old gray whales reported by Blokhin (1990; 1997) and a single gray whale cow with a calf photographed off Alaska (8.33m) on 6 July 1998 (Perryman unpublished data), gray whales are estimated to be between 8.0 and 8.5m when weaned. We believe, as suggested by Sumich (1986), that growth in young gray whales slows at weaning and that the length of yearling gray whales average 8.5m, as originally suggested by Rice and Wolman (1971). The estimate of the mean length of yearling gray whales is based on the interpretation of the overall length distribution of photographed whales and must be taken with caution until a larger sample of lengths for whales of known age is available.

Gray whales do not feed significantly during their winter migration. Rice and Wolman (1971) estimated that gray whales passing San Francisco on their way north weigh about 11-29% less on average than they did passing the same site on their way south. A detectable pattern of reduction was found in the relationship between maximum width and length that mirrors the pattern of change in nutritive condition reported from specimens. Southbound parturient females were found to be widest relative to their length and northbound lactating females were the narrowest. For whales that were not parturient or with a calf, southbound gray whales were significantly wider than those migrating north. These findings indicate that the documented changes in overall nutritive condition in gray whales associated with fasting for approximately 60 days can be detected in measurements from photographs. It was also found that the location of the widest point on a gray whale changes with reproductive and nutritive condition. These results imply that, for stranded gray whales and those taken in fisheries, axillary girth might be a less informative measurement than maximum girth coupled with the distance from the rostral tip to the site of this measurement. Published results on both fin and sei whales also indicate that girth measurements taken near mid-length or even farther posterior may be more reflective of changes in condition than those taken at the apex of the pectoral fin (Lockyer *et al.*, 1985; Lockyer, 1987; Víkingsson, 1990).

ACKNOWLEDGEMENTS

This research was supported by funding provided by the National Marine Fisheries Service, Office of Protected Species. Our research would not be possible without the cameras and maintenance support provided by the US Department of Defense. We are especially indebted to the support provided by Mrs. Loris Jacobs of the US Army's Material Command for her tireless efforts to track down reconnaissance assets that were due for retirement. We also thank Jim Gilpatrick for his support on survey flights and

John Brandon for helping with calculation of median calving dates. This paper was improved significantly by reviews and comments by Robert Brownell (SWFSC), Douglas DeMaster (NMML), Meghan Donahue (SWFSC), Steve Reilly (SWFSC) and David Rugh (NMML).

REFERENCES

Angliss, R.P., Rugh, D.J., Withrow, D.E. and Hobbs, R.C. 1995. Evaluations of aerial photogrammetric length measurements of the Bering-Chukchi-Beaufort Seas stock of bowhead whales (*Balaena mysticetus*). *Rep. int. Whal. Commn* 45:313-24.

Atchley, W.R., Gaskins, C.T. and Anderson, D. 1976. Statistical properties of ratios. I. Empirical results. *Syst. Zool.* 25:563-83.

Best, P.B. and Rüther, H. 1992. Aerial photogrammetry of southern right whales, *Eubalaena australis*. *J. Zool., London.* 228:595-614.

Blokhin, S.A. 1984. Investigations of gray whales taken in the Chukotka coastal waters, USSR. pp. 487-509. *In:* M.L. Jones, S.L. Swartz and S. Leatherwood (eds.) *The Gray Whale,* Eschrichtius robustus. Academic Press Inc., Orlando, Florida. xxiv+600pp.

Blokhin, S.A. 1990. Results of investigation of the California-Chukota population of gray whales taken off Chukotka in 1980-1989. Paper SC/A90/G26 presented to the special meeting of the Scientific Committee on the Assessment of Gray whales, Seattle, April 1990 (unpublished). 26pp. [Paper available from the Office of this Journal].

Blokhin, S.A. 1997. The results of studies of the American population of gray whales taken in the coastal waters of the Chukotka Peninsula in 1996. Paper SC/49/AS15 presented to the IWC Scientific Committee, Bournemouth, September 1997 (unpublished). [Paper available from the Office of this Journal].

Blokhin, S.A. and Tiupeleev, P.A. 1987. Morphological study of the earplugs of gray whales and the possibility of their use in age determination. *Rep. int. Whal. Commn* 37:341-5.

Bogoslovskaya, L.S., Votrogov, L.M. and Semenova, T.N. 1981. Feeding habits of the gray whale off Chukotka. *Rep. int. Whal. Commn* 31:507-10.

Bryant, P.J., Lafferty, C.M. and Lafferty, S.K. 1984. Reoccupation of Laguna Guerrero Negro, Baja California, Mexico, by gray whales. pp. 375-87. *In:* M.L. Jones, S.L. Swartz and S. Leatherwood (eds.) *The Gray Whale,* Eschrichtius robustus. Academic Press, Inc., Orlando, Florida. xxiv+600pp.

Buckland, S.T. and Breiwick, J.M. 2002. Estimated trends in abundance of eastern Pacific gray whales from shore counts (1967/68 to 1995/96). *J. Cetacean Res. Manage.* 4(1):In press.

Day, R.W. and Quinn, G.P. 1989. Comparisons of treatments after an analysis of variance in ecology. *Ecol. Monogr.* 59:433-63.

Gilpatrick, J.W. 1996. Calibration of radio altimeter readings used in aerial photogrammetry of eastern tropical Pacific dolphins 1992-1993. NOAA Tech. Mem. NMFS-SWFSC-226. 11pp. [Available from: *http://www.nmfs.gov*].

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,* Eschrichtius robustus. Academic Press Inc., Orlando, Florida. xxiv+600pp.

Jones, M.L. and Swartz, S.L. 1984. Demography and phenology of gray whales and evaluation of whale-watching activities in Laguna San Ignacio, Baja California Sur, Mexico. pp. 309-74. *In:* M.L. Jones, S.L. Swartz and S. Leatherwood (eds.) *The Gray Whale,* Eschrichtius robustus. Academic Press, Inc., Orlando, Florida. xxiv+600pp.

Koski, W.R., Davis, R.A., Miller, G.W. and Withrow, D.E. 1992. Growth rates of bowhead whales as determined from low-level aerial photogrammetry. *Rep. int. Whal. Commn* 42:491-9.

Lockyer, C. 1987. The relationship between body fat, food resource and reproductive energy costs in North Atlantic fin whales (*Balaenoptera physalus*). *Symp. Zool. Soc., Lond.* 57:343-61.

Lockyer, C.H., McConnell, L.C. and Waters, T.D. 1985. Body condition in terms of anatomical and biochemical assessments of body fat in North Atlantic fin and sei whales. *Can. J. Zool.* 63:2,328-38.

Norris, K.S. and Gentry, R.L. 1974. Capture and harnessing of young California gray whales, *Eschrichtius robustus*. *Mar. Fish. Rev.* 36(4):58-64.

Pacheco, J.A.S. 1998. Gray whale mortality at Ojo de Liebre and Guerro Negro lagoons, Baja California Sur, Mexico: 1984-1995. *Mar. Mammal Sci.* 14:155-61.

Perryman, W.L. and Lynn, M.S. 1993. Identification of geographic forms of common dolphin (*Delphinus delphis*) from aerial photogrammetry. *Mar. Mammal Sci.* 9(2):119-37.

Perryman, W.L. and Westlake, R.L. 1998. A new geographic form of the spinner dolphin, *Stenella longirostris*, detected with aerial photogrammetry. *Mar. Mammal Sci.* 14(1):38-50.

Pike, G.C. 1962. Migration and feeding of the gray whale (*Eschrichtius gibbosus*). *J. Fish. Res. Bd Can.* 19(5):815-38.

Poole, M.M. 1984. Migration corridors of gray whales along the central California coast, 1980-1982. pp. 389-407. *In:* M.L. Jones, S.L. Swartz and S. Leatherwood (eds.) *The Gray Whale,* Eschrichtius robustus. Academic Press, Inc., Orlando, Florida. xxiv+600pp.

Read, A.J. 1990. Estimation of body condition in harbor porpoises, *Phocoena phocoena*. *Can. J. Zool.* 68(9):1962-6.

Reist, J.D. 1985. An empirical evaluation of several univariate methods that adjust for size variation in morphometric data. *Can. J. Zool.* 63:1429-39.

Rice, D.W. 1983. Gestation period and fetal growth of the gray whale. *Rep. int. Whal. Commn* 33:539-44.

Rice, D.W. 1990. Life history parameters of the gray whale: a review of published estimates. Paper SC/A90/G22 presented to the special meeting of the Scientific Committee on the Assessment of Gray whales, Seattle, April 1990 (unpublished). 6pp. [Paper available from the Office of this Journal].

Rice, D.W. and Wolman, A.A. 1971. *The Life History and Ecology of the Gray Whale* (Eschrichtius robustus). American Society of Mammalogists, Special Publication No. 3, Stillwater, Oklahoma. viii+142pp.

Rice, D.W., Wolman, A.A., Withrow, D.E. and Fleischer, L.A. 1981. Gray whales on the winter grounds in Baja California. *Rep. int. Whal. Commn* 31:477-93.

Rugh, D.J., Shelden, K.E.W. and Schulman-Janiger, A. 2001. Timing of the gray whale southbound migration. *J. Cetacean Res. Manage.* 3(1):31-9.

Shelden, K.E.W., Rugh, D.J. and Boeve, S.A. 1997. Gray whale calf sightings collected by the National Marine Mammal Laboratory during southbound migrations, 1952-95. Paper SC/47/AS4 presented to the IWC Scientific Committee, May 1995, Dublin (unpublished). [Paper available from the Office of this Journal].

Sumich, J.L. 1986. Growth in young gray whales (*Eschrichtius robustus*). *Mar. Mammal Sci.* 2(2):145-52.

Sumich, J.L., Goff, T. and Perryman, W.L. 2001. Growth of two captive gray whale calves. *Aquat. Mamm.* 27(3):231-3.

Víkingsson, G.A. 1990. Energetic studies on fin and sei whales caught off Iceland. *Rep. int. Whal. Commn* 40:365-73.

Yablokov, A.V. and Bogoslovskaya, L.S. 1984. A review of Russian research on the biology and commercial whaling of the gray whale. pp. 465-85. *In:* M.L. Jones, S.L. Swartz and S. Leatherwood (eds.) *The Gray Whale,* Eschrichtius robustus. Academic Press Inc., Orlando, Florida. xxiv+600pp.

Zimushko, V.V. and Ivashin, M.V. 1980. Some results of Soviet investigations and whaling of Gray whales (*Eschrichtius robustus*, Lilljeborg, 1961). *Rep. int. Whal. Commn* 30:237-46.

A framework for evaluating *Strike Limit Algorithms* for populations reduced to small numbers

ANDRÉ E. PUNT* and JEFFREY M. BREIWICK+

Contact e-mail: aepunt@u.washington.edu

ABSTRACT

A generic framework is outlined within which operating models for populations reduced to low numbers but still subject to exploitation ('type 3 fisheries') can be developed. This framework is founded on an individual-based operating model that includes temporally correlated environmental variation in births and survival as well as the possibility of occasional catastrophic reductions in survival. Methods are developed to specify the value of the parameter that determines the productivity of the resource from that for *MSYR*, to enable simulation trials based on this framework to be parameterised in terms of *MSYR*. Three potential candidate *Strike Limit Algorithms* are evaluated using 14 'generic' simulation trials that capture a range of factors pertinent to type 3 fishery situations. The 'Maximum-likelihood-like' *SLA* developed for use in the management of the Bering-Chukchi-Beaufort Seas stock of bowhead whales performs adequately for many of these 14 trials, but not all. In contrast, a variant of the 'PBR approach' is shown to perform adequately in terms of achieving conservation objectives for all of the trials. The information needed to specify trials for actual type 3 fishery situations is outlined.

KEYWORDS: ABORGINAL WHALING; INDIVIDUAL-BASED MODEL; STRIKE LIMIT ALGORITHM; MODELLING

INTRODUCTION

The Scientific Committee of the International Whaling Commission (IWC) has, since 1996, been developing candidate *Strike Limit Algorithms*[1] (*SLA*s) for the management of aboriginal subsistence whaling. The approach adopted by the Scientific Committee to contrast the performances of candidate *SLA*s is Monte Carlo simulation. This approach was used to develop the Revised Management Procedure (RMP) for commercial whaling (Kirkwood, 1997) and has also been applied to develop management procedures for several fish stocks (see, for example, the review by Butterworth and Punt (1999)).

IWC (1997) determined that (initial) priority should be given to developing *SLA*s for two types of fishery:

(1) 'type 1 fishery': a case where there is relatively little available information and stock identity problems and where the Scientific Committee has had considerable problems in providing management advice;
(2) 'type 2 fishery': a case where there is a relatively large amount of information and satisfactory management advice can be provided.

Subsequently (IWC, 1998), a third fishery type was identified. This type is characterised by a high extent of depletion to a small (of the order of 300 animals) total population size. Progress towards the selection of an *SLA* for one of the type 2 fisheries (the Bering-Chukchi-Beaufort Seas stock of bowhead whales) has been completed (IWC, 2003) and simulation trials have been designed for one example of a type 1 fishery (IWC, 1998). In contrast, relatively little progress has been made towards selecting *SLA*s for type 3 fisheries (although see Breiwick and DeMaster, 1999).

What distinguishes type 3 situations (and consequently may necessitate a qualitatively different approach to their management and the associated evaluations) is the relatively much greater influence on their dynamics of 'demographic stochasticity' (annual variation in the number of births and deaths as a consequence of the fact that the population is comprised of individual animals), environmental stochasticity (annual variation in birth and survival rates common to groups of individuals) and catastrophic events (infrequent events, but ones that have a large impact on groups of individuals)[2] (Shaffer, 1981; Lande, 1993).

The objectives of management for aboriginal whaling have been set by the IWC as follows:

(1) ensure risks of extinction not seriously increased (highest priority);
(2) enable harvests in perpetuity appropriate to cultural and nutritional requirements (implicit in this is the concept of catch stability);
(3) maintain stocks at highest net recruitment level and if below that ensure they move towards it.

Arguably, the most important step in the process of evaluating candidate *SLA*s is the development of an appropriate set of simulation trials. This set reflects different assumptions for an operating model[3]. The trials for fishery types 1 and 2 have been largely based on density-dependent age- and sex-structured operating models, characterised by the BALEEN II model underlying the HITTER-FITTER package (de la Mare, 1989; Punt, 1999). Although some of the operating models for the type 1 and 2 fisheries allow for demographic stochasticity and environmental stochasticity in births (e.g. IWC, 2000), they are 'lumped' (in the sense that the behaviour of individuals is ignored and all animals of a given age are assumed to be interchangeable). It is unclear, however, whether such lumped operating models are appropriate for evaluating *Strike Limit Algorithms* for type 3

[1] An algorithm that determines the number of allowable strikes from a population subject to an aboriginal harvest.

[2] The impact of catastrophic events may be more severe for small than for large populations because the population may be reduced to levels at which demographic stochasticity may render the resource extinct.

[3] The model that represents the true situation for the simulation trials.

* *School of Aquatic and Fishery Sciences, Box 355020, University of Washington, Seattle, WA 98105-5020, USA.*
+ *National Marine Mammal Laboratory, 7600 Sand Point Way NE, Bldg. 4, Seattle, WA 98115, USA.*

fisheries because the behaviour of individuals may play a large role in the dynamics of the population when the population size is small.

This paper therefore develops an operating model that is individual-based. The behaviour of this model as the values for its parameters are changed is examined and this behaviour is contrasted qualitatively with that of a lumped operating model. The performances of three potential candidate *SLA*s (including one based on the PBR approach (Wade, 1998) and another based on the 'Maximum-likelihood-like' *SLA* developed by Punt (2001)) are then contrasted by using a set of illustrative fishery type 3 trials. Finally, the further work needed to apply the framework developed in this paper to actual type 3 fisheries (e.g. Cook Inlet white whales (*Delphinapterus leucus*), Baffin Bay bowhead whales (*Balaena mysticetus*) and the humpback whale (*Megaptera novaeangliae*) feeding aggregation off West Greenland) is identified.

METHODS

The stochastic individual-based model

The model described below considers the population at the level of the individual. The information available for each animal includes: sex; age; whether recruited to the fishery or not; whether mature or not (females only - animals aged 60 and older are assumed not to calve); the animal's mother and the year in which the animal (if female and mature) last calved. The following steps occur during each year y to project the model forward.

(1) Generate the survival rate for year y for animals of sex s and age a, $S^s_{y,a}$ (the survival probability can be temporally correlated or the population can be subject to catastrophic mortality events - see Appendix A for details).
(2) Compute the number of 1+, mature and recruited animals at the start of year y.
(3) Compute the probability of a mature animal giving birth, f_y (see Appendix B).
(4) Determine, for each mature female that did not give birth the previous year or whose calf died during its first year, whether it gives birth at the start of year y by conducting a Bernoulli trial with probability f_y.
(5) Determine the sex of each new calf by conducting a Bernoulli trial with probability 0.5 for each calf.
(6) Remove the catch. If fishing involves a fixed harvest rate, a Bernoulli trial with a probability of success equal to the harvest rate is conducted for each recruited animal to determine whether it is caught or not. If the numbers caught (by sex) rather than the fishing mortality rate is specified, then this catch is removed by selecting animals at random (and without replacement) from the set of recruited animals (i.e. multinomially). If a female that calved during year y is harvested during year y, its calf is assumed to die.
(7) Determine, for each animal that survives the harvest, whether it survives natural mortality by conducting a Bernoulli trial with probability $S^s_{y,a}$. If a female that calved during year y dies due to natural causes during year y, its calf is assumed to die as well.
(8) Determine, for each unrecruited animal, whether it recruits at the end of the year by conducting a Bernoulli trial with probability of success of

$$\delta_a = (R_a - R_{a-1})/(1 - R_{a-1})$$

where R_a) is the probability that an animal of age a is recruited[4].
(9) Determine, for each immature female, whether it matures at the end of the year by conducting a Bernoulli trial with probability $(M_a - M_{a-1})/(1 - M_{a-1})$ where M_a is the probability that an animal of age a is mature.
(10) Increment the age of each animal.

Specifying the values for the parameters of the model

The values for the biological parameters of the model have been set to those 'typical' for a large baleen whale (based on the suggestions by Breiwick and DeMaster (1999) - Table 1). For consistency with previous *SLA* evaluations (IWC, 2002), all animals are assumed to recruit to the fishery at age 1, i.e. $R_0 = 0$ and $R_1 = 1$ for ages $a = 1$ and older. The value of the parameter z has been assumed to be 2.39. This choice for z corresponds to *Maximum Sustainable Yield Level* ($MSYL$) = 0.6 for a model in which the recruitment and maturity ogives are identical and density-dependence acts on the mature (or recruited) component of the population.

The remaining parameters of the model determine the current (1+) population size, the current depletion (current 1+ abundance relative to its average pre-exploitation level), $MSYR_{1+}$[5], the impact of catastrophic events, whether survival is subject to temporally correlated fluctuations and whether the probability of birth is subject to temporally correlated fluctuations. Fourteen scenarios (Table 2) examine the sensitivity of the results to a range of hypotheses about these factors. Current (1+) abundance is low (300) for all but two cases and the base-case value for *MSYR* is 2.5%. The majority of the trials ignore catastrophic events and temporally correlated birth and survival rates (equivalent to the SD (Demographic Stochasicity) assumptions underlying the trials developed for the Bering-Chukchi-Beaufort Seas stock of bowhead whales - IWC, 2002).

The specifications related to productivity in Table 2 are expressed in terms of the *MSY* rate, *MSYR*. However, the model parameter determining productivity is A and not *MSYR* (see Equation B.1). Unlike the deterministic case

[4] This equation arises from the relationship between δ_a and R_a: $R_a+1 = R_a + \delta_a(1 - R_a)$, i.e. the number of animals recruiting at age a+1 is δ_{a+1} multiplied by the number of animals of age a+1 that have not yet recruited.
[5] *MSYR* is the ratio of *MSY* to the population size at which *MSY* is achieved.

Table 1
Expected survival probabilities and the probability of being mature as a function of age.

Age	0	1	2	3	4	5	6	7	8	9	10+
Survival, S_a	0.75	0.75	0.8	0.85	0.9	0.92	0.94	0.95	0.96	0.97	0.98
Maturity, M_a	0	0	0.25	0.375	0.5	0.5	0.625	0.75	1	1	1

Table 2
Specifications of the 14 simulation trials.

Description	Trial no.	Current population size	Current depletion	Catastrophic events	*CV*, ρ of env. variation in survival	CV, $\tilde{\rho}$ of env. variation in births	$MSYR_{1+}$ (%)
Base-case	1a	300	0.05	No	0	0	2.5
	1b	300	0.1	No	0	0	2.5
Low $MSYR_{1+}$	2b	300	0.05	No	0	0	1
	2b	300	0.1	No	0	0	1
High $MSYR_{1+}$	3a	300	0.05	No	0	0	4
	3b	300	0.1	No	0	0	4
Large population size	4a	3,000	0.05	No	0	0	2.5
	4b	3,000	0.1	No	0	0	2.5
With catastrophic events	5a	300	0.05	Yes*	0	0	2.5
	5b	300	0.1	Yes*	0	0	2.5
With env. variation in survival	6a	300	0.05	No	0.025, 0.71	0	2.5
	6b	300	0.1	No	0.025, 0.71	0	2.5
With env. variation in births	7a	300	0.05	No	0	0.5, 0.71	2.5
	7b	300	0.1	No	0	0.5, 0.71	2.5

* A catastrophic event involves an 80% reduction in population size The probability of a catastrophic event is 2% each year.

(Punt, 1996), the expected sustainable catch is not related analytically to *A* and *MSYR*. Instead, it is necessary to apply a numerical approach to solve for *A* given *MSYR* and the values for the remaining parameters of the model (Appendix C). This approach is similar to that used in New Zealand to define *MSY* for fish species with highly variable recruitment (Francis, 1992). The initial conditions for the population projections correspond to a resource at its (deterministic) pre-exploitation level. Two levels for the depletion at the start of the simulations (0.05 and 0.1) are considered. Both of these correspond to a resource well below conventional target levels.

Each simulation trial involves projecting the population from pre-exploitation equilibrium for 200 years without catches (so that the age-structure at the start of the first year in which historical catches are taken is not in equilibrium because of the impact of random variation in births and deaths), and then removing 100 years of historical catches (the catches are removed under the assumption of a constant intended harvest rate over the 100 years – the actual harvest rate (the ratio of the number harvested to the 1+ population size) will differ from the intended harvest rate given the Bernoulli process used to decide whether an animal is harvested or not). The 250 simulations that constitute a simulation trial are each constructed by selecting the random variates that determine stochasticity in birth and death rates and in catastrophic events and then varying the harvest rate over the historical period so that the depletion at the end of this period equals the pre-specified depletion. This implies that the historical catches (which are, of course, integer numbers) differ among simulations.

Evaluating *Strike Limit Algorithms*

Three candidate *SLA*s are evaluated for the 14 scenarios:

(1) A constant catch strategy - this strategy sets the strike limit each year to the level of (integer) catch which achieves a median final depletion for trial 1 as close as possible to that for the 'Maximum-likelihood-like' *SLA* below.
(2) The Potential Biological Removals (PBR) strategy (see Appendix D for a brief overview).
(3) The 'Maximum-likelihood-like' *SLA* - a variant of an *SLA* considered for the Bering-Chukchi-Beaufort Seas stock of bowhead whales (see Appendix E for a brief overview).

Each simulation trial involves 250 simulations of a 100-year projection period in which the strike limit is set every fifth year. The data available to the *SLA*s are (unbiased) estimates of 1+ abundance (and their CVs) generated using the protocol applied for Bering-Chukchi-Beaufort Seas stock of bowhead whales (see Section B.1 of Appendix 3 of IWC (2002)). These abundance estimates are generated every fifth year with a CV (if the population was 60% of its pre-exploitation equilibrium level) of 0.25, starting the year before the *SLA* is first applied (i.e. it is assumed that no historical estimates of abundance are available). Additional variance in abundance estimates is ignored for the purposes of these trials. The level of need is taken to be infinite for these trials and the historical annual catches supplied to the *SLA* are taken to be the average (over simulations) catch each year (truncated to the nearest integer).

RESULTS AND DISCUSSION

Basic model features

Simulations with no environmental variability or catastrophic events

Fig. 1 plots expected (i.e. average over years and simulations) catch versus exploitation rate and expected catch versus expected number of recruited (equals 1+) animals for trials 1b, 2b and 3b (pre-exploitation size for the 1+ component of the population, K^{1+}, of 3,000 and $MSYR_{1+}$ rates of 2.5, 1 and 4% respectively). These figures were constructed by projecting an unexploited population forward for 800 years under a variety of levels of intended constant exploitation rate. The results for the last 400 years of the 800-year period for 500 such projections were then averaged to obtain the results for each exploitation rate plotted in Fig. 1. The yield curves in Fig. 1 are sufficiently smooth that it can be concluded that the number of replicates conducted (500) was (more than) adequate to determine the value of *A* reasonably accurately. $MSYL_{1+}$ does not occur at $0.6K^{1+}$ for any of these trials (53-56% for the results reported in Fig. 1). This is, however, to be expected given that the recruitment and maturity ogives differ substantially (Punt, 1996). The ratio of the exploitation rate at which the population is rendered extinct within 800 years to $MSYR_{1+}$ decreases as $MSYR_{1+}$ is increased. This ratio is 4 for $MSYR_{1+} = 1\%$ but less than 2 for $MSYR_{1+} = 4\%$.

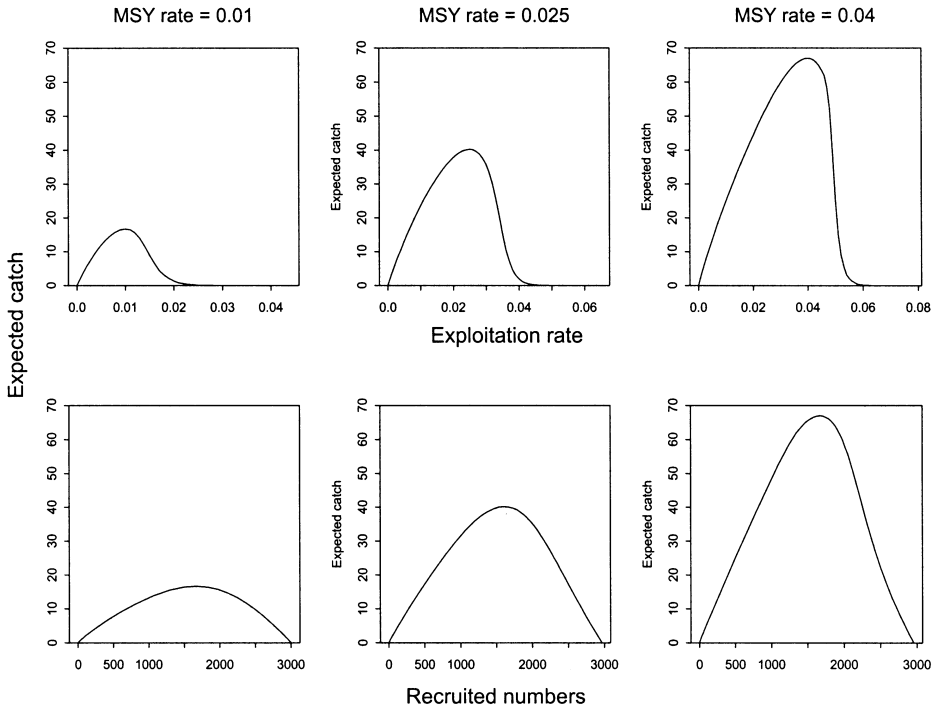


Fig. 1. Expected catch versus exploitation rate and expected number of recruited animals for trials in which $K^{1+} = 3{,}000$ and $MSYR_{1+}$ = 1, 2.5, and 4%.

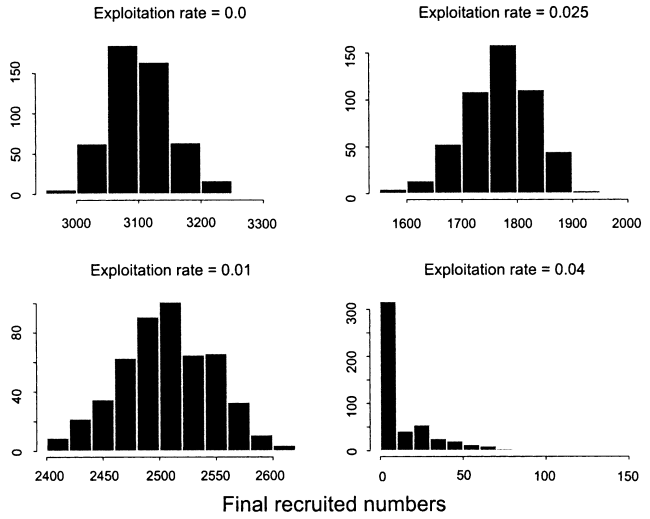


Fig. 2. Distribution of the number of recruited animals after 800 years of the application of a (intended) constant exploitation rate. $MSYR_{1+} = 2.5\%$ and $K^{1+} = 3{,}000$ for this figure.

Fig. 2 plots the distribution of population size (500 points) after 800 years for trial 1b ($MSYR_{1+} = 2.5\%$; $K^{1+} = 3{,}000$; population initially at K^{1+}) for fixed exploitation rates of 0, 1, 2.5 and 4% while Fig. 3 shows the time-trajectories of 1+ population size for the first 10 (of 500) simulations for each of these exploitation rates for this trial.

As expected (e.g. May *et al.*, 1978; Sissenwine *et al.*, 1988), Fig. 2 indicates that the variability in the distributions of the number of recruited animals after 800 years increases with increasing exploitation rate (CVs for the final population size of 1.6% for exploitation rates of 0 and 1%, 3.4% for an exploitation rate of 2.5%, and > 100% for an exploitation rate of 4%).

The time-trajectories of 1+ population size reach approximate equilibrium after about 100 years, except when the exploitation rate is 4% when the resource is predicted to be rendered extinct eventually (Fig. 3). It should be noted that although there is variability among simulations (Fig. 2), this variability is not particularly large in comparison to the impact of the harvest (Fig. 3).

Simulations with environmental variability and catastrophic events

Fig. 4 shows 800-year time-trajectories of 1+ population size for the base-case population model (trial 1b) and variants thereof in which allowance is made for catastrophic events, (correlated) environmental variation in survival, and (correlated) environmental variation in births (trials 1b, 5b, 6b and 7b). The value of K^{1+} equals 3,000, $MSYR_{1+} = 2.5\%$, the exploitation rate is 2.5% and the population is initially at K^{1+} for all of the results reported in Fig. 4.

The factor that has the largest impact (i.e. the largest difference from the base-case) is environmental variation in the survival rate. In contrast, the one factor included in the current trials for fishery type 2 (environmental variation in births) has relatively the lowest of the impacts in Fig. 4 (except, of course, for the base-case which does not include any sources of environmental variation). It is noteworthy that when there is environmental variation in survival, the resource is almost rendered extinct on occasion even though the harvest rate equals *MSYR*. The fact that environmental variation in survival has the largest impact in Fig. 4 may appear surprising given the arguments of Mangel and Tier (1994) and Gerber and Hilborn (2001) that catastrophic events are often the most important determinants of the probability of population extinction. However, this can be attributed to the specific choices for the values for the parameters that determine environmental variation of survival and the impact (and frequency) of catastrophic events (for which there is little basis in data). The fact that environmental variation in births has the least impact is probably a fairly general result because environmental variation in survival and catastrophes impact several age-classes whereas the impact of environmental variation in births is restricted to calves.

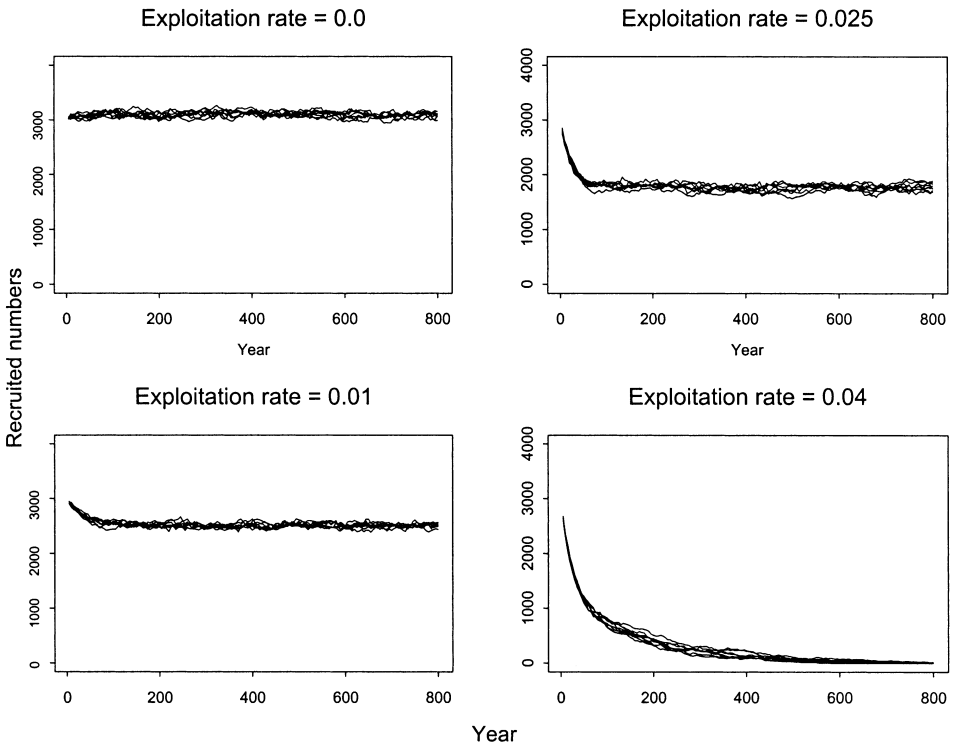


Fig. 3. Time-trajectories of 1+ population size for ten simulations for four exploitation rates when $K^{1+} = 3{,}000$ and $MSYR_{1+} = 2.5$.

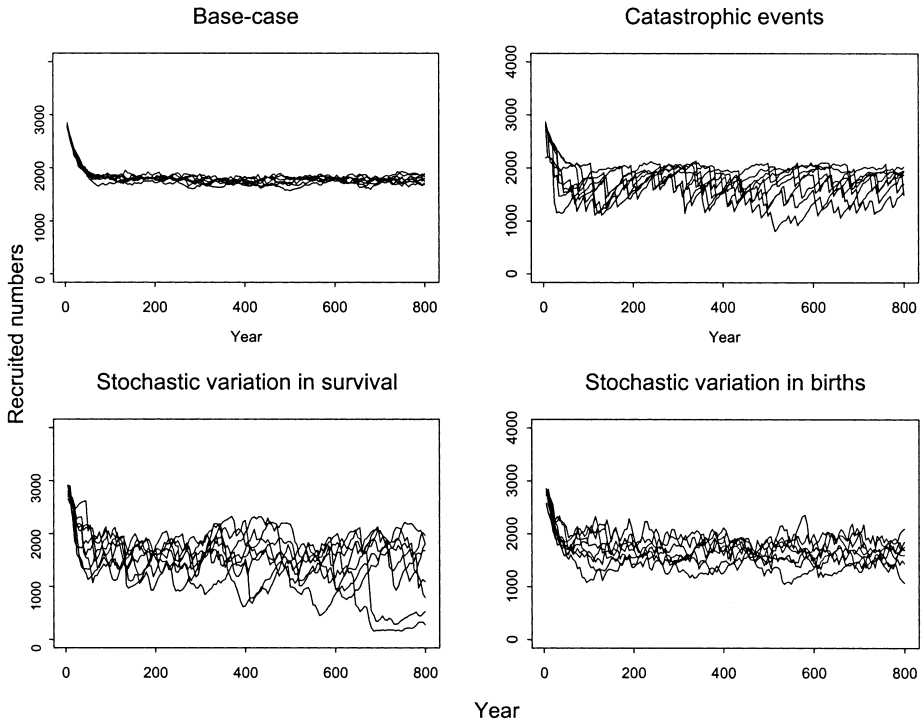


Fig. 4. Ten time-trajectories of 1+ population size under an exploitation rate of 2.5% based on different amounts (and types) of environmental variability when $K^{1+} = 3{,}000$ and $MSYR_{1+} = 2.5$.

Fig. 5 contrasts the implications of different levels for the coefficient of variation (CV) of the probability of survival and the extent of inter-annual correlation in annual survival. The 1+ population size is 300 animals (10% of K^{1+}) at the start of the 100-year projection period and no catches are taken for the entire 100 years. The range for the CV of survival is bounded above by 0.1 because, given the specifications related to survival (Table 1), there is no solution to Equation (A.3) for CVs larger than 0.1. The variability in the trajectories increases with inter-annual correlation in survival and with the CV for the probability of survival. In 250 simulations over 100 years, no extinctions occurred for CVs < 0.075 but several extinctions occurred for CVs of 0.075 and larger. The population size after 100 years substantially exceeds K^{1+} in some instances and the frequency of this increases with increasing CV. This occurs because, when the CV of survival is high and given that survival is bounded above by 1, the median probability of survival exceeds the mean probability of survival by an amount that increases as the CV increases. Therefore,

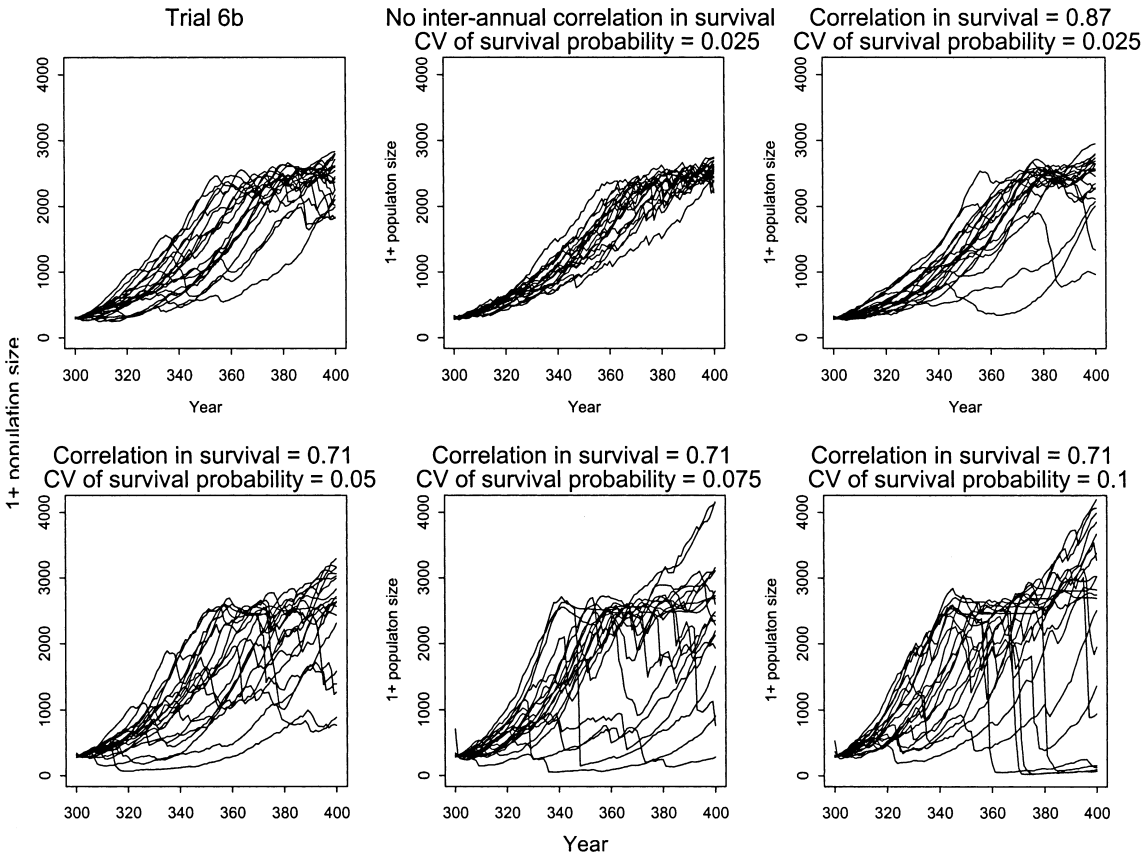


Fig. 5. Twenty 100-year time-trajectories of 1+ population size under zero harvest for various specifications related to the extent of environmental variation in survival for the scenario in which the population is depleted to 300 animals (10% of K^{1+}) at the start of year 300.

although the expected population size in the absence of exploitation is K^{1+}, the higher the variability in the probability of survival, the more substantially the population will exceed K^{1+} on occasion.

Comparison with a lumped model

A deterministic lumped model cannot mimic the types of behaviour evident in Figs 4 and 5 for the model variants in which there is environmental variability in survival or in births. In contrast, the level of variability due solely to 'demographic stochasticity' (see Fig. 3 and top left panel of Fig. 4) is sufficiently small that the results of a deterministic model would not be qualitatively different.

It would be difficult to implement some of the features of the individual-based model within a standard lumped model (for example, the assumption that if a calf dies, its mother is more likely to calve immediately, and, particularly, that the catch is taken multinomially from the population). An additional feature that would be difficult to implement using a standard lumped model would be the regulation that mothers with calves cannot be harvested. In addition, the use of an individual-based model as the operating model ensures that the number of animals (for each age and sex) is always an integer; the lumped model developed by IWC (2002) does not ensure that this is the case.

The individual-based model framework is therefore clearly more flexible in terms of the assumptions that it can represent. However, this flexibility is not without (computational) cost. In particular, the software required to implement the individual-based model is substantially more complex than that to implement the corresponding lumped model, has markedly larger storage requirements (it is necessary to store several pieces of information for each animal) and runs considerably more slowly. The impact of the last two disadvantages increases with the number of animals. For example, the time required to evaluate *SLA*s using operating models 4a and 4b is at least an order of magnitude longer than is the case for the other operating models. However, for type 3 fisheries, the population size is relatively small by definition, so that these computational constraints are not as prohibitive as might be the case if the individual-based model formed the basis for an operating model for, say, the eastern North Pacific stock of gray whales for which the current population size is over 20,000.

While it is clear that deterministic lumped models are inappropriate as the basis for operating models for type 3 fisheries, it may well be case that models that pool individuals of a given age and maturity state (could give birth this year, gave birth last year, etc.) and allow for demographic and environmental stochasticity in births and deaths could mimic the behaviour of an individual-based model adequately. Evaluation of this issue is, however, beyond the scope of the current paper but should form a focus for future work.

Evaluation of *SLA*s

Table 3 contrasts the performances of the three *SLA*s described above and the strategy of setting zero strike limits, in terms of a subset of the performance statistics used to evaluate *SLA*s for other aboriginal whaling operations (e.g. IWC, 2002). The performance statistics are the 5th, median and 95th points of the distributions for the following quantities:

Table 3

Performance statistics for four alternative *SLA*s for the 14 simulation trials.

(a) Zero catches for 100 years.

Description	Trial no.	Final depletion (1+)			Final depletion (mat)			Lowest depletion (1+)			Relative recovery (1+)			Relative recovery (mat)		
		5%	Median	95%	5%	Median	95%	5%	Median	95%	5%	Median	95%	5%	Median	95%
Base-case	1a	0.752	0.799	0.825	0.899	1.001	1.058	0.045	0.049	0.054	14.7	16.0	17.5	16.3	18.4	21.2
	1b	0.816	0.845	0.871	0.993	1.035	1.077	0.095	0.100	0.104	8.0	8.5	8.9	8.5	9.5	11.1
Low $MSYR_{1+}$	2a	0.149	0.200	0.256	0.160	0.222	0.290	0.045	0.050	0.054	3.0	4.0	5.0	3.1	4.0	5.3
	2b	0.302	0.392	0.505	0.322	0.435	0.592	0.095	0.099	0.104	3.0	3.9	5.1	3.0	4.0	5.1
High $MSYR_{1+}$	3a	0.788	0.819	0.846	0.987	1.010	1.035	0.046	0.050	0.054	15.1	16.5	17.9	17.3	19.9	23.2
	3b	0.865	0.901	0.943	0.941	0.974	1.009	0.095	0.100	0.104	8.5	9.0	9.6	8.5	9.5	10.9
Large population size	4a	0.790	0.804	0.813	0.978	1.011	1.033	0.046	0.051	0.054	14.9	15.8	17.1	17.2	18.2	19.7
	4b	0.836	0.846	0.856	1.021	1.033	1.045	0.096	0.101	0.104	8.1	8.4	8.8	9.0	9.5	10.1
With catastrophic events	5a	0.572	0.809	0.843	0.650	1.015	1.101	0.043	0.049	0.054	11.8	15.8	18.0	12.9	18.3	22.0
	5b	0.671	0.842	0.908	0.775	1.044	1.102	0.087	0.099	0.104	6.7	8.4	9.2	7.0	9.4	10.8
With env. variation in survival	6a	0.355	0.789	0.878	0.401	0.978	1.158	0.044	0.049	0.054	7.4	15.5	18.3	7.7	17.7	23.1
	6b	0.585	0.840	0.942	0.624	0.997	1.160	0.086	0.099	0.104	6.0	8.5	9.5	6.0	9.3	11.4
With env. variation in births	7a	0.664	0.777	0.889	0.787	0.975	1.103	0.045	0.050	0.054	13.1	15.4	18.0	14.0	17.7	20.9
	7b	0.738	0.828	0.940	0.897	1.010	1.124	0.095	0.099	0.104	7.3	8.3	9.5	7.7	9.3	10.9

(b) The 'Maximum likelihood-like' *SLA*.

Description	Trial no.	Final depletion (1+)			Final depletion (mat)			Lowest depletion (1+)			Relative recovery (1+)			Relative recovery (mat)			Average annual catch			AAV		
		5%	Median	95%	5%	Median	95%	5%	Median	95%	5%	Median	95%	5%	Median	95%	5%	Median	95%	5%	Median	95%
Base-case	1a	0.074	0.558	0.808	0.077	0.641	1.028	0.037	0.049	0.054	1.6	11.2	17.1	1.8	11.8	19.6	0.2	10.0	21.4	0.039	0.084	0.720
	1b	0.385	0.690	0.868	0.445	0.805	1.060	0.095	0.100	0.104	4.0	6.8	8.8	4.1	7.5	10.4	0.2	11.9	19.1	0.033	0.053	0.800
Low $MSYR_{1+}$	2a	0.011	0.177	0.235	0.013	0.197	0.268	0.004	0.049	0.054	0.2	3.6	4.7	0.2	3.6	4.8	0.2	0.3	5.2	0.103	0.322	0.800
	2b	0.168	0.348	0.477	0.189	0.382	0.547	0.092	0.098	0.104	1.7	3.5	4.8	1.7	3.5	4.7	0.1	0.2	4.2	0.077	0.350	0.800
High $MSYR_{1+}$	3a	0.017	0.716	0.828	0.018	0.811	1.028	0.013	0.049	0.054	0.3	14.3	17.5	0.4	15.8	22.4	0.3	40.3	57.4	0.029	0.041	0.771
	3b	0.622	0.716	0.903	0.662	0.804	0.997	0.095	0.099	0.104	6.2	7.2	9.3	6.4	7.8	10.0	0.3	32.5	39.0	0.025	0.034	0.633
Large population size	4a	0.071	0.226	0.780	0.076	0.249	0.966	0.032	0.049	0.054	1.4	4.4	15.8	1.4	4.5	17.9	11.5	139.4	203.3	0.033	0.047	0.211
	4b	0.364	0.625	0.833	0.402	0.719	1.035	0.096	0.101	0.104	3.6	6.2	8.0	3.7	6.6	9.2	4.2	139.3	172.2	0.034	0.044	0.362
With catastrophic events	5a	0.039	0.595	0.843	0.042	0.687	1.093	0.009	0.051	0.089	1.4	10.4	22.6	1.5	11.7	25.2	0.2	0.5	30.1	0.037	0.240	0.760
	5b	0.272	0.701	0.876	0.292	0.794	1.091	0.042	0.096	0.186	2.1	6.3	15.5	2.1	6.7	18.2	0.2	10.4	25.4	0.033	0.060	0.743
With env. variation in survival	6a	0.036	0.547	0.859	0.039	0.619	1.132	0.005	0.045	0.155	0.9	9.1	30.6	1.0	9.8	33.8	0.2	0.6	45.2	0.037	0.171	0.829
	6b	0.185	0.671	0.882	0.205	0.772	1.126	0.021	0.101	0.282	1.7	5.6	16.8	1.7	6.1	21.6	0.2	7.6	28.6	0.031	0.077	0.850
With env. variation in births	7a	0.018	0.612	0.843	0.019	0.715	1.093	0.006	0.048	0.082	0.6	11.9	20.2	0.6	13.0	22.4	0.2	0.7	27.4	0.038	0.216	0.771
	7b	0.332	0.664	0.901	0.390	0.783	1.075	0.057	0.102	0.160	3.2	6.1	12.2	3.1	6.7	14.0	0.2	11.2	20.8	0.035	0.055	0.880

(c) Constant catch strategy

Description	Trial no.	Final depletion (1+)			Final depletion (mat)			Lowest depletion (1+)			Relative recovery (1+)			Relative recovery (mat)			Average annual catch		
		5%	Median	95%	5%	Median	95%	5%	Median	95%	5%	Median	95%	5%	Median	95%	5%	Median	95%
Base-case	1a	0.477	0.652	0.749	0.531	0.756	0.918	0.045	0.049	0.054	10.0	13.1	14.8	10.2	14.0	16.9	5.0	5.0	5.0
	1b	0.725	0.790	0.820	0.861	1.003	1.052	0.094	0.099	0.104	7.3	7.9	8.4	8.1	9.1	10.3	5.0	5.0	5.0
Low $MSYR_{1+}$	2a	0.003	0.042	0.090	0.004	0.054	0.113	0.003	0.039	0.052	0.1	0.8	1.8	0.1	1.0	2.0	4.9	5.0	5.0
	2b	0.003	0.078	0.182	0.007	0.101	0.227	0.003	0.073	0.100	0.0	0.8	1.8	0.1	0.9	1.9	4.8	5.0	5.0
High $MSYR_{1+}$	3a	0.761	0.786	0.814	0.996	1.017	1.037	0.045	0.050	0.054	14.5	15.8	17.2	17.5	20.0	23.3	5.0	5.0	5.0
	3b	0.792	0.836	0.876	0.949	0.987	1.025	0.095	0.099	0.104	7.9	8.4	8.8	8.5	9.7	11.0	5.0	5.0	5.0
Large population size	4a	0.780	0.799	0.810	0.957	0.999	1.022	0.046	0.051	0.054	14.8	15.7	16.9	17.0	18.1	19.4	5.0	5.0	5.0
	4b	0.832	0.841	0.851	1.024	1.035	1.048	0.096	0.101	0.104	8.1	8.4	8.8	9.0	9.5	10.1	5.0	5.0	5.0
With catastrophic events	5a	0.031	0.623	0.827	0.044	0.734	1.082	0.017	0.051	0.089	1.6	10.6	16.7	1.8	11.9	18.6	5.0	5.0	5.0
	5b	0.000	0.754	0.882	0.000	0.913	1.078	0.000	0.096	0.186	0.0	6.6	10.0	0.0	7.4	11.9	4.1	5.0	5.0
With env. variation in survival	6a	0.000	0.596	0.874	0.000	0.685	1.139	0.000	0.045	0.151	0.0	8.1	19.5	0.0	9.1	22.9	1.4	5.0	5.0
	6b	0.000	0.744	0.946	0.000	0.862	1.134	0.000	0.099	0.282	0.0	5.1	11.6	0.0	5.5	14.4	1.2	5.0	5.0
With env. variation in births	7a	0.106	0.630	0.851	0.121	0.747	1.049	0.027	0.049	0.082	3.5	10.9	16.1	3.7	12.3	17.3	5.0	5.0	5.0
	7b	0.316	0.762	0.888	0.362	0.950	1.091	0.056	0.102	0.157	4.2	7.0	9.5	4.7	8.0	11.0	5.0	5.0	5.0

(d) PBR approach

Description	Trial no.	Final depletion (1+)			Final depletion (mat)			Lowest depletion (1+)			Relative recovery (1+)			Relative recovery (mat)			Average annual catch			AAV		
		5%	Median	95%	5%	Median	95%	5%	Median	95%	5%	Median	95%	5%	Median	95%	5%	Median	95%	5%	Median	95%
Base-case	1a	0.499	0.603	0.697	0.561	0.693	0.832	0.045	0.049	0.054	10.2	12.2	13.8	10.5	12.8	15.4	7.7	9.9	12.8	0.070	0.103	0.140
	1b	0.712	0.759	0.784	0.821	0.924	0.984	0.095	0.100	0.104	7.1	7.6	8.0	7.5	8.5	9.6	7.4	8.9	10.5	0.063	0.088	0.117
Low $MSYR_{1+}$	2a	0.084	0.117	0.156	0.090	0.132	0.182	0.044	0.049	0.054	1.7	2.3	3.2	1.6	2.4	3.4	1.9	2.8	3.6	0.133	0.185	0.265
	2b	0.170	0.231	0.309	0.191	0.266	0.360	0.093	0.098	0.104	1.7	2.3	3.1	1.8	2.4	3.2	2.2	3.0	4.0	0.097	0.138	0.188
High $MSYR_{1+}$	3a	0.759	0.783	0.808	0.951	0.979	1.006	0.045	0.050	0.054	14.4	15.8	17.2	16.8	19.3	22.5	16.4	19.2	22.8	0.056	0.080	0.102
	3b	0.785	0.816	0.852	0.927	0.963	1.004	0.095	0.099	0.104	7.7	8.2	8.7	8.4	9.4	10.7	11.2	12.4	14.0	0.059	0.078	0.102
Large population size	4a	0.499	0.577	0.630	0.559	0.656	0.728	0.046	0.051	0.054	10.0	11.3	12.4	10.4	12.0	13.2	86.2	101.1	115.5	0.072	0.095	0.133
	4b	0.719	0.749	0.770	0.852	0.899	0.937	0.096	0.101	0.104	7.1	7.5	7.9	7.7	8.3	8.8	78.0	87.6	98.0	0.062	0.082	0.108
With catastrophic events	5a	0.356	0.621	0.760	0.385	0.715	0.940	0.043	0.049	0.054	7.2	12.4	15.4	7.2	13.0	17.5	6.0	10.5	14.7	0.068	0.097	0.152
	5b	0.517	0.751	0.817	0.572	0.904	1.021	0.085	0.099	0.104	5.1	7.5	8.3	5.5	8.2	9.8	5.7	8.6	11.0	0.064	0.089	0.124
With env. variation in survival	6a	0.224	0.618	0.819	0.253	0.717	1.074	0.042	0.049	0.054	4.7	12.2	16.9	4.7	13.0	20.2	4.5	10.1	17.5	0.067	0.106	0.170
	6b	0.393	0.750	0.846	0.450	0.889	1.103	0.081	0.098	0.104	4.0	7.5	8.6	4.2	8.2	10.6	4.3	9.2	11.9	0.060	0.088	0.134
With env. variation in births	7a	0.429	0.591	0.730	0.477	0.680	0.855	0.045	0.050	0.054	8.7	11.8	14.6	8.9	12.4	15.9	7.2	10.0	13.3	0.069	0.097	0.143
	7b	0.609	0.724	0.846	0.707	0.890	1.020	0.094	0.099	0.104	6.2	7.2	8.6	6.3	8.1	9.7	6.5	8.5	10.4	0.064	0.087	0.118

(1) the final depletion - the ratio of the population size (1+ and mature) at the end of the projection period to the average pre-exploitation level.
(2) the lowest depletion – the ratio of the lowest 1+ population size during the projection period to the average pre-exploitation level.
(3) the relative recovery – the ratio of the population size (1+ and mature) at the end of the projection period to that at the start of this period.
(4) the average annual catch.
(5) the average absolute variation (AAV) in annual catch.

A variety of other performance statistics are used by the IWC Scientific Committee to evaluate *SLA*s. However, these statistics are not currently easily amenable to simulation trials based on individual-based models.

The values for the performance statistics all indicate that, in the absence of future catches, the population size after 100 years will exceed that at the start of the projection period (Table 3a). The extent of increase (as measured by the 'relative recovery' statistics) differs depending on the *MSYR* and the initial state of the resource (5 or 10% of the average pre-exploitation level). The final depletions are, as expected, lower if the strike limits are set using the 'Maximum-likelihood-like' *SLA* (Table 3b). However, some recovery is guaranteed in many of the cases. The exceptions to this (all of which have an initial depletion of 0.05) are trials 2a (initial depletion = 0.05; $MSYR_{1+}$ = 1%), 3a (initial depletion = 0.05; $MSYR_{1+}$ = 4%), 6a (initial depletion = 0.05; environmental variation in survival), and 7a (initial depletion = 0.05; environmental variation in births) for which the lower 5th percentiles for the relative recovery statistic are less than unity. The reasons for the poor performance for trials 2a, 6a and 7a are readily apparent: these are cases in which productivity is poor or survival is subject to environmental stochasticity. The poor performance for trial 3a is surprising; this is a case in which productivity is high, and substantial resource recovery occurs in the bulk of the simulations. It would seem to be the consequence of a few large catch limits in the first years of the 100-year period leading to a high exploitation rate at low population size.

The value of applying a feedback harvest strategy can be evaluated by comparing the results in Tables 3b and 3c. The constant level of catch for Table 3c was chosen so that the median final depletion for trial 1 was as close as is possible (given that the strike limit has to be an integer) to that for the 'Maximum-likelihood-like' *SLA*. This level of constant catch is 5 and, in fact, the median final depletion for trial 1 substantially exceeds that for 'Maximum-likelihood-like' *SLA*. However, the performance of the constant catch *SLA* is very poor (for example, extinction occurs in more than 5% of simulations for trials 5b, 6a and 6b and more than 50% of the final depletions for trials 2a and 2b are below the initial depletions), which indicates that there is considerable value in applying a feedback harvest strategy for type 3 fisheries.

The PBR approach (Table 3d) is inherently more conservative than the 'Maximum-likelihood-like' *SLA* (note the generally lower average catches/strikes). Consequently, this *SLA* does not lead to values for the lower 5th percentile of the relative recovery statistic less than unity for any of the trials. More importantly, the PBR approach does not drop the resource appreciably below the level that would be achieved in the absence of any catches for trials 5a-7b (contrast the lower 5thpercentiles of the lowest depletion distributions in Tables 3a and 3d).

Worked needed to tailor the framework to specific type 3 fisheries

The information required to develop case-specific trials for actual type 3 fisheries includes the following:

(1) historical catches.
(2) biological parameters (survival rates, fecundity rates, minimum calving interval).
(3) extent of environmental variability in survival and birth rate.
(4) probability and expected severity of catastrophic events.
(5) nature and frequency of future data collection programmes.
(6) relative (age- and sex-specific) probability of being harvested.
(7) scenarios regarding *MSYR* and *MSYL*.
(8) the bounds on the levels of need that an *SLA* will have to be able to cope with.
(9) performance statistics.

The information for type 3 fisheries tends to be even worse than is the case for type 1 and type 2 fisheries. This may mean that instead of being able to specify values for some of the parameters of the operating model exactly, ranges for the values for these parameters will have to be considered. One implication of this is that the range of scenarios that may need to be considered may be very wide (to ensure that the actual situation is covered by the set of trials). The extent of environmental variation in birth and survival rates and the probability and severity of catastrophic events cannot be estimated for any type 3 fisheries. However, it is possible that inferences from other whale species (e.g. Angliss *et al.*, 1995; Perryman *et al.*, 1997) could be used to develop ranges for, for example, the extent of environmental variation in births. Developing scenarios regarding catastrophic events using data for cetaceans will be even more problematic than developing such scenarios for the extent of environmental variation in birth and survival rates. Gerber and Hilborn (2001) provide data related to catastrophic events for a range of otariid species and it may be possible to utilise information of this type when developing scenarios.

The primary data available for type 1 and 2 fisheries on which strike limits might be based has been assumed to be unbiased estimates of total (1+) abundance. Unfortunately, the absolute abundance data for some type 3 fisheries are likely to comprise only very infrequent (and possibly highly biased) estimates of population size. For such cases, it may be necessary to use the information on relative abundance and from the age-structure of the catch to achieve reasonable *SLA* performance.

The scenarios for *MSYR* could be expressed in terms of the expected rate of recovery at very low stock size. Values for this rate could plausibly be inferred from data for other whale species (e.g. Best (1993)).

The specification of a need envelope (i.e. bounds on possible levels of need) would have to be achieved through discussions at the Commission, with technical advice from the Scientific Committee. The selection of appropriate tunings of candidate *SLA*s would need to be coordinated between the IWC Scientific Committee and the Commission.

General discussion

The framework in this paper could be extended substantially depending on circumstances. There are several areas where additional model development work may be required for

some type 3 fisheries, for example: Allee effects; time-dependence in catastrophic events; catastrophic impacts on the birth rate; changes over time in the environment and spatial structure. The first of these has been ignored in the analyses of this paper because, although there must be a 'level below which recovery is impossible', there are too few data to define this level generically. In fact, there may be some good reasons not to define this level but (as was the case during the development of the RMP) to instead interpret performance on the basis of the lowest depletion statistic. The individual-based model can be extended to include spatial structure (as has been done for lumped models for the *Implementation Simulation Trials* developed for the North Atlantic, Southern Hemisphere and North Pacific minke whales). However, as was the case for the RMP, it would seem advisable to first examine the generic properties of *SLA*s for type 3 fisheries before submitting them to evaluation against detailed spatially structured case-specific trials.

The process of conditioning case-specific trials may be substantially more complicated if operating models for type 3 fisheries are to be founded on individual-based models. For example, the approach used to condition the trials for the Bering-Chukchi-Beaufort Seas stock of bowhead whales and of those for the eastern North Pacific stock of gray whales involves placing a prior on the current population size. Unfortunately, it is not easy to 'solve' an individual-based model for K such that the current abundance (or depletion) equals a particular value. In this paper, the constraint that the depletion of the 1+ component of the population after the period of historical catches equals 0.05 or 0.1 was achieved subject to a tolerance of 0.005. This aspect of the conditioning problem could be overcome by placing a prior on K rather than on current depletion or current abundance, although this approach to conducting Bayesian assessments has been criticised in the past in the Scientific Committee (e.g. Butterworth and Punt, 1995).

ACKNOWLEDGEMENTS

Support for this work was provided by NOAA/NMFS. Doug DeMaster and Anne York (National Marine Mammal Laboratory, Seattle) are thanked for valuable discussions and suggestions. Doug Butterworth is thanked for his extensive comments on an earlier version of this paper. Geof Givens made the suggestion of using the logistic transformation to deal with environmental variation in birth rate.

REFERENCES

Angliss, R.P., Rugh, D.J., Withrow, D.E. and Hobbs, R.C. 1995. Evaluations of aerial photogrammetric length measurements of the Bering-Chukchi-Beaufort Seas stock of bowhead whales (*Balaena mysticetus*). *Rep. int. Whal. Commn* 45:313-24.

Best, P.B. 1993. Increase rates in severely depleted stocks of baleen whales. *ICES J. Mar. Sci.* 50(3):169-86.

Breiwick, J.M. and DeMaster, D.P. 1999. Exploratory type 3 fishery simulations. Paper SC/51/AWMP8 presented to the IWC Scientific Committee, May 1999, Grenada, WI (unpublished). 9pp. [Paper available from the Office of this Journal].

Butterworth, D.S. and Punt, A.E. 1995. On the Bayesian approach suggested for the assessment of the Bering-Chukchi-Beaufort Seas stock of bowhead whales. *Rep. int. Whal. Commn* 45:303-11.

Butterworth, D.S. and Punt, A.E. 1999. Experiences in the evaluation and implementation of management procedures. *ICES J. Mar. Sci.* 56(6):985-98.

de la Mare, W.K. 1989. Report of the Scientific Committee, Annex L. The model used in the HITTER and FITTER programs (Program:FITTER.SC40). *Rep. int. Whal. Commn* 39:150-1.

Francis, R.I.C.C. 1992. Recommendations concerning the calculation of maximum constant yield (MCY) and current yield (CAY). NZ Fisheries Research Assessment Document 92/8.

Gerber, L.R. and Hilborn, R. 2001. Catastrophic events and recovery from low densities in populations of otariids: implication for risk of extinction. *Mammal Rev.* 11:131-50.

International Whaling Commission. 1997. Report of the Scientific Committee. Annex I. Report of the Workshop on the Development of an Aboriginal Subsistence Whaling Management Procedure (AWMP). *Rep. int. Whal. Commn* 47:192-202.

International Whaling Commission. 1998. Report of the Scientific Committee. Annex I. Report of the Workshop on the Development of an Aboriginal Subsistence Whaling Management Procedure (AWMP). *Rep. int. Whal. Commn* 48:203-36.

International Whaling Commission. 2000. Report of the Scientific Committee. Annex E. Report of the Standing Working Group on the Development of an Aboriginal Subsistence Whaling Management Procedure. *J. Cetacean Res. Manage. (Suppl.)* 2:125-54.

International Whaling Commission. 2002. Report of the Third Workshop on the Development of an Aboriginal Subsistence Whaling Management Procedure (AWMP). Annex F. Examination of a simplified version of the Inertia model. *J. Cetacean Res. Manage. (Suppl.)* 4:450-3.

International Whaling Commission. 2003. Report of the Scientific Committee. *J. Cetacean Res. Manage. (Suppl.)* 5:In press.

Kirkwood, G.P. 1997. The Revised Management Procedure of the International Whaling Commission. pp. 91-9. *In:* E.K. Pikitch, D.D. Huppert and M.P. Sissenwine (eds.) *Global Trends: Fisheries Management.* American Fisheries Society Symposium 20, Bethseda, Maryland. xv+328pp.

Lande, R. 1993. Risks of population extinction from demographic and environmental stochasticity and random catastrophes. *Am. Nat.* 142:911-27.

Mangel, M. and Tier, C. 1994. Four facts every conservation biologist should know about persistence. *Ecology* 75:607-14.

May, R.M., Beddington, J.R., Horwood, J.W. and Shepherd, J.G. 1978. Exploiting natural populations in an uncertain world. *Math. Biosci.* 42:219-52.

Perryman, W.L., Donahue, M.A., Reilly, S.B. and Perkins, P.C. 1997. Annual calf production for the California stock of gray whales, 1994-1997. Paper SC/49/AS13 presented to the IWC Scientific Committee, September 1997, Bournemouth (unpublished). 26pp. [Paper available from the Office of this Journal].

Punt, A.E. 1996. The effects of assuming that density dependence in the HITTER-FITTER model acts on natural mortality rather than fecundity. *Rep. int. Whal. Commn* 46:629-36.

Punt, A.E. 1999. Report of the Scientific Committee. Annex R. A full description of the standard BALEEN II model and some variants thereof. *J. Cetacean Res. Manage. (Suppl.)* 1:267-76.

Punt, A.E. 2001. A revised Maximum-likelihood-like' Strike Limit Algorithm for the Bering-Chukchi-Beaufort Seas stock of bowhead whales. Paper SC/53/AWMP1 presented to the IWC Scientific Committee, July 2001, London (unpublished). [Paper available from the Office of this Journal].

Shaffer, M.L. 1981. Minimum population sizes for species conservation. *Biosci.* 31:131-4.

Sissenwine, M.P., Fogarty, M.J. and Overholtz, W.J. 1988. Some fisheries management implications of recruitment variability. pp. 129-52. *In:* J.A. Gulland (ed.) *Fish Population Dynamics.* 2nd. Edn. John Wiley and Sons, Chichester. xvii+422pp.

Wade, P.R. 1998. Calculating limits to the allowable human-caused mortality of cetaceans and pinnipeds. *Mar. Mammal Sci.* 14:1-37.

Appendix A

GENERATING THE AGE- AND SEX-SPECIFIC SURVIVAL RATES

The probability that an animal of sex *s* and age *a* survives the impact of natural mortality, $S^s_{y,a}$, when subject to temporally correlated environmental variation and to occasional 'catastrophic' reductions in survival, is given by:

$$S^s_{y,a} = \frac{e^{q^s_{y,a}}}{e^{q^s_{y,a}}+1} U_y \quad q^s_{y,a} = \mu^s_a + \rho(q^s_{y-1,a} - \mu^s_a) + \varepsilon_y \sigma^s_a \qquad \text{(A.1)}$$

where:

$S^s_{y,a}$ is the survival probability for animals of sex *s* and age *a* during year *y*;
$q^s_{y,a}$ is the (logit-transformed) survival probability for animals of sex *s* and age *a* during year *y* in the absence of catastrophic events;
μ^s_a is the mean of the distribution for $q^s_{y,a}$;
ρ is the temporal correlation in the environmental anomalies in survival rate;
σ^s_a is the standard deviation of the survival anomalies;
ε_y is a random variate generated from $N(0;1^2)$;
U_y is a factor to account for (temporally independent) 'catastrophic' reductions in survival rate:

$$U_y = \begin{cases} U_{low} & \text{if } \tilde{\varepsilon}_y < p_{low} \\ (1 - U_{low}\, p_{low})/(1 - p_{low}) & \text{otherwise} \end{cases} \qquad \text{(A.2)}$$

U_{low} is the fraction by which survival is reduced if a 'catastrophic' event takes place;
p_{low} is the probability of a 'catastrophic' event taking place; and
ε_y is a random variate generated from U[0,1].

The values for the μ^s_as and σ^s_as are chosen so that the expected survival rate and the coefficient of variation of survival rate equal pre-specified values — $E(S^s_a)$ and $CV(S^s_a)$ respectively. This involves choosing values for μ^s_a and σ^s_a to satisfy the following system of equations:

$$E(S^s_a) = \int_{-\infty}^{\infty} \frac{e^q}{1+e^q} \frac{\sqrt{1-\rho^2}}{\sqrt{2\pi}\,\sigma^s_a} e^{-\frac{1-\rho^2}{2(\sigma^s_a)^2}(q-\mu^s_a)^2} dq$$

$$CV(S^s_a) = \frac{1}{E(S^s_a)} \sqrt{\int_{-\infty}^{\infty} \left(\frac{e^q}{1+e^q}\right)^2 \frac{\sqrt{1-\rho^2}}{\sqrt{2\pi}\,\sigma^s_a} e^{-\frac{1-\rho^2}{2(\sigma^s_a)^2}(q-\mu^s_a)^2} dq - E(S^s_a)^2} \qquad \text{(A.3)}$$

The formulation for U_y is selected so that the expected value of U_y equals 1.

Appendix B

DETERMINING THE BIRTH RATE PROBABILITY

The expected probability during year *y* that a mature female that has not given birth for at least *x* years gives birth, $E(f_y)$ is given by:

$$E(f_y) = f_0(1 + A(1 - (P^{1+}_y/P^{1+}_{-\infty})^z)) \qquad \text{(B.1)}$$

where:

f_0 is the birth rate at the average pre-exploitation level;
A is the 'resilience' parameter;
z is the 'degree of compensation' parameter;
P^{1+}_y is the number of 1+ animals at the start of year *y*; and
$P^{1+}_{-\infty}$ is the (average) number of 1+ animals in an unexploited state.

The value of f_0 is given by:

$$f_0 = \frac{B_{-\infty}/N^m_{-\infty}}{1 - \frac{1}{N^m_{-\infty}} \sum_{y=2}^{x} B_{-\infty}\, e^{-M(y-2)}} \qquad \text{(B.2)}$$

where:

$B_{-\infty}$ is the (average) number of births (calves) in an unexploited state;
x is the minimum calving interval;
M is the instantaneous rate of natural mortality on mature animals; and
$N^m_{-\infty}$ is the (average) number of mature animals in an unexploited state.

The realised value for the probability during year *y* of a mature female that has not given birth for at least x years giving birth is generated using the equation:

$$f_y = \frac{e^{\hat{q}_y}}{e^{\hat{q}_y}+1} \qquad \hat{q}_y = \mu_y + \tilde{\rho}(\hat{q}_{y-1} - \mu_{y-1}) + \eta_y \sigma_F \qquad \text{(B.3)}$$

where:

μ_y is selected (G. Givens, pers. comm.) so that:

$$E(f_y) = \int_{-\infty}^{\infty} \frac{e^q}{1+e^q} \frac{\sqrt{1-\tilde{\rho}^2}}{\sqrt{2\pi}\,\sigma_F} e^{-\frac{1-\tilde{\rho}^2}{2(\sigma_F)^2}(q-\mu_y)^2} dq \qquad \text{(B.4)}$$

σ_F is a measure of the environmental variation in births;
$\tilde{\rho}$ is the temporal correlation in birth rate; and
η_y is a random variate generated from $N(0;1^2)$.

Appendix C

SOLVING FOR *A* GIVEN *MSYR*

MSYR can be defined by the equation:

$$\left.\frac{dC}{dF}\right|_{F=MSYR} = 0 \qquad \text{(C.1)}$$

where $C(F)$ is the catch (expected catch for a stochastic model) as a function of the intended exploitation rate F.

A numerical approach for computing A given *MSYR* is therefore to search (using, for example, a bisection method) for the value of A such that

$$C(MSYR + \Delta F) - C(MSYR - \Delta F) = 0.$$

This is a straightforward calculation for deterministic models for which the function $C(F)$ is well defined (e.g. Punt (1996; 1999)). Unfortunately, this is not the case for stochastic models. For such models, it is necessary to define $C(F)$ as the average of the catch when the exploitation rate is fixed at F. In this paper, $C(F)$ for a single replicate is obtained by projecting from the average pre-exploitation level for 800 years fixing the exploitation rate to F. The average catch over the last 400 years of the 800-year projection period is then taken to be $C(F)$ for that replicate. The choice of 400 years for the 'burn in' period was selected based on projections for the base-case assumptions under a range of values for the intended exploitation rate, F.

Appendix D

THE PBR APPROACH

The strike limit for year y, Q_y, is defined according to the equation:

$$Q_y = F_R\, 0.5\, R_{\max}\, N_{\min}(y) \qquad \text{(D.1)}$$

where:

F_R is the 'recovery factor' – assumed to be 0.5 for the analyses of this paper;
$R_{\max}$ is the maximum theoretical net productivity rate (assumed to be 0.04 – the default value for cetaceans – Wade, 1998); and
$N_{\min}(y)$ is a 'minimum population size' (the lower 20th% confidence limit) based on the most recent estimate of abundance.

Appendix E

THE 'MAXIMUM-LIKELIHOOD' *SLA*

The population dynamics model and likelihood function that form the basis for this *SLA* are:

$$N_{y+1} = N_y + 1.4184\, MSYR\, N_y(1 - (N_y/K)^2) - C_y \quad \text{(E.1)}$$

$$L = \prod_{y<y^\bullet} \exp\left(-\frac{1}{2}\frac{\ell n(N_y / N_y^{obs})^2}{(CV_y^{obs})^2 + \{CV_{add}(y - y^\bullet)/10\}^2}\right) \qquad \text{(E.2)}$$

where:

N_y is the number of (1+) animals (of both sexes) at the start of year y;
K is the pre-exploitation number of 1+ animals;
MSYR is the *MSY* rate parameter;
C_y is the catch (both sexes) during year y;
$y^\bullet$ is the year for which a strike limit is required;
N_y^{obs} is the estimate of 1+ abundance for year y;
CV_y^{obs} is the observed coefficient of variation for N_y^{obs}; and
CV_{add} is the 'additional' coefficient of variation – defined as the CV added to that for surveys conducted ten years before the year for which a strike limit is required.

The calculation of a strike limit for year $y^\bullet$ involves first calculating a 'raw' strike limit for year $y^\bullet$, $Q_{y^\bullet}^R$, and modifying this to conform with constraints on inter-5-year-block variability in strike limits. The 'raw' strike limit for year $y^\bullet$ is computed using the following algorithm.

(1) Find, for 50 values for *MSYR* spaced equally between 0 and 0.05, the values for K that maximise Equation (E.2).
(2) Assign weights to each of the 50 combinations of *MSYR* and K obtained at step (1) using the formula $\exp(-likelihood)\, P(MSYR)$ where *likelihood* is the negative log-likelihood and the prior is of the form:

[6] This population dynamics model is a modification of a discrete logistic model with the maximum rate of increase parameterised in terms of *MSYR*.

$$P(MSYR) = \begin{cases} P_{\max} & \text{if } MSYR \le 0.01 \\ P_{\max} + \frac{1-P_{\max}}{0.02}(MSYR - 0.01) & \text{if } 0.01 < MSYR < 0.03 \\ 1 & \text{if } MSYR \ge 0.03 \end{cases} \quad \text{(E.3)}$$

where $P_{\max}$ is the maximum of the 'prior' for *MSYR*.

These 50 combinations of *MSYR* and *K* with their corresponding weights (normalised to sum to 1) form the 'posterior' distribution for steps (3) and (4).

(3) Find the strike limit that satisfies the criterion that Q_1^{th} 'posterior' percentile of $P_{y\cdot+20}/K$ equals a pre-specified value $\tilde{MSYL}$ (the value for $P_{y\cdot+20}/K$ is computed by projecting the model 20 years into the future assuming that catch is constant from year $y^{\bullet}$ to year $y^{\bullet}$ +20).

(4) Find the strike limit that satisfies the criterion that the Q_2^{th} posterior percentile of $P_{y\cdot+20}/P_{y\cdot}$ equals a pre-specified value $Z_{2,crit}$.

The constraints on inter-block variability in strike limits are imposed using the formula:

$$P(MSYR) = \begin{cases} P_{\max} & \text{if } MSYR \le 0.01 \\ P_{\max} + \frac{1-P_{\max}}{0.02}(MSYR - 0.01) & \text{if } 0.01 < MSYR < 0.03 \\ 1 & \text{if } MSYR \ge 0.03 \end{cases} \quad \text{(E.4)}$$

where:

$Q_{y\cdot}^{\bullet}$ is the strike limit for year $y^{\bullet}$ following modification to constrain variability in strike limits; and

β_1, β_2 are the parameters that determine the constraints on inter-block variation in catch limits.

The values for the tuning parameters are taken to be $P_{\max} = 5$, $CV_{add} = 0.025$, $\tilde{MSYL} = 0.6$, $Z_{2,crit} = 1.1$, $Q_1 = 0.19$, $Q_2 = 0.19$, $\beta_1 = 0.25$, and $\beta_2 = 1.75$.

A blue whale (*Balaenoptera musculus*) feeding ground in a southern Australian coastal upwelling zone

PETER C. GILL

School of Ecology and Environment, Deakin University and Australocetus Research, C/- Post Office, Narrawong Victoria 3285, Australia
Contact e-mail: pcgill@ozemail.com.au

ABSTRACT

A localised aggregation of blue whales, which may be pygmy blue whales (*B. m. brevicauda*), occurs in southern Australian coastal waters (between 139°45′E-143°E) during summer and autumn (December-May), where they feed on coastal krill (*Nyctiphanes australis*), a species which often forms surface swarms. While the abundance of blue whales using this area is unknown, up to 32 blue whales have been sighted in individual aerial surveys. Krill appear to aggregate in response to enhanced productivity resulting from the summer-autumn wind-forced Bonney Coast upwelling along the continental shelf. During the upwelling's quiescent (winter-spring) period, blue whales appear to be absent from the region. Krill surface swarms have been associated with 48% of 261 blue whale sightings since 1998, with direct evidence of feeding observed in 36% of all sightings. Mean blue whale group size was 1.55 (SD = 0.839), with all size classes represented including calves. This seasonally predictable upwelling system is evidently a regular feeding ground for blue whales, and careful management of human activities is required there.

KEYWORDS: BLUE WHALE; AUSTRALASIA; FEEDING GROUNDS; EUPHAUSIIDS; OCEANOGRAPHY; HABITAT

INTRODUCTION

Blue whales (*Balaenoptera musculus*) in the Southern Hemisphere were so severely hunted during the 20th century (Clapham *et al.*, 1999) that sightings are rare (Branch and Butterworth, 2001). Under Australia's Federal environment legislation they are listed as a Critically Endangered Threatened Species, and are the subject of a forthcoming Recovery Plan (Commonwealth of Australia, 1999). Feeding areas outside Antarctic waters are almost unknown. Most available information about blue whale feeding ecology comes from the Northern Hemisphere, where feeding grounds have been identified in coastal upwelling zones off the coast of southern California (Croll *et al.*, 1998; 2000; Fiedler *et al.*, 1998), Baja California (Reilly and Thayer, 1990; D. Gendron, pers. comm.), and in the Gulf of Saint Lawrence, Canada (Simard *et al.*, 1986; Sears *et al.*, 1990). Blue whales also feed in offshore upwelling areas in the eastern tropical and equatorial Pacific (Reilly and Thayer, 1990; Palacios, 1999). In all these areas, blue whales feed on abundant (in patches) euphausiid prey.

Historically, blue whales have rarely been sighted in Australian waters (Bannister *et al.*, 1996). During December 1995, a dedicated Blue Whale Cruise (under the auspices of the International Whaling Commission) was conducted through southern Australian waters, from Fremantle, Western Australia, to Hobart, Tasmania. During the cruise, an aggregation of blue whales was found off Discovery Bay near the Victoria-South Australia border (141°E). Many were identified by experienced observers as pygmy blue whales (*B.m. brevicauda*; see description in Ichihara, 1966). Some feeding behaviour and defaecation was observed, and euphausiids were collected from surface swarms for future identification (Kato *et al.*, 1996).

Historical data for western Victoria show some strandings and sightings records of blue whales since 1887 (Anon., 1999). Furthermore, the cold-water summer-autumn Bonney Coast upwelling has been known to exist in this area (e.g. Schahinger, 1987), suggesting that the presence of feeding blue whales might be related to this persistent and predictable oceanographic event. The Bonney Coast upwelling is a rare feature along the southern mainland coast of Australia, which is generally characterised as a nutrient-poor region fed by the warm water of the Leeuwin Current (e.g. Herzfeld, 1997). A yacht-based survey in February 1998 established that blue whales were again present in this region (author's data). Blue whales were found surface feeding in Discovery Bay and a longer-term ecological study was initiated, focusing on the linkages between blue whales and this upwelling habitat.

MATERIALS AND METHODS

Study area

The study initially focused on Discovery Bay, but was later expanded to cover the entire continental shelf from Robe, South Australia (139°45′E), to Port Campbell, Victoria (143°E), an area of approximately 12,000km^2 (Fig. 1). The east-west limits of the feeding area have not yet been conclusively delineated.

Survey methods

Aerial surveys

Surveys began in December 1998 and continued until May 2001. Most intensive coverage occurred during the summer-autumn upwelling seasons, with some additional coverage during winter and spring. Aerial surveys were designed to provide a measure of seasonal occurrence and general distribution, relative to environmental features such as the upwelling and bathymetry, and therefore were not intended as a method of establishing absolute or relative abundance of blue whales. Survey tracks covered much of the continental shelf and to at least 5 n.miles (9.25km) to seaward of it; flying further to seaward was precluded by cost and safety considerations. Complete coverage of the entire feeding area was not possible in any one flight, with flight frequency and tracklines determined by funding constraints, changing weather conditions, research needs and perceived changes in whale distribution with time. Spatial analysis of survey effort, whale distribution and biological oceanography will be addressed in subsequent publications.

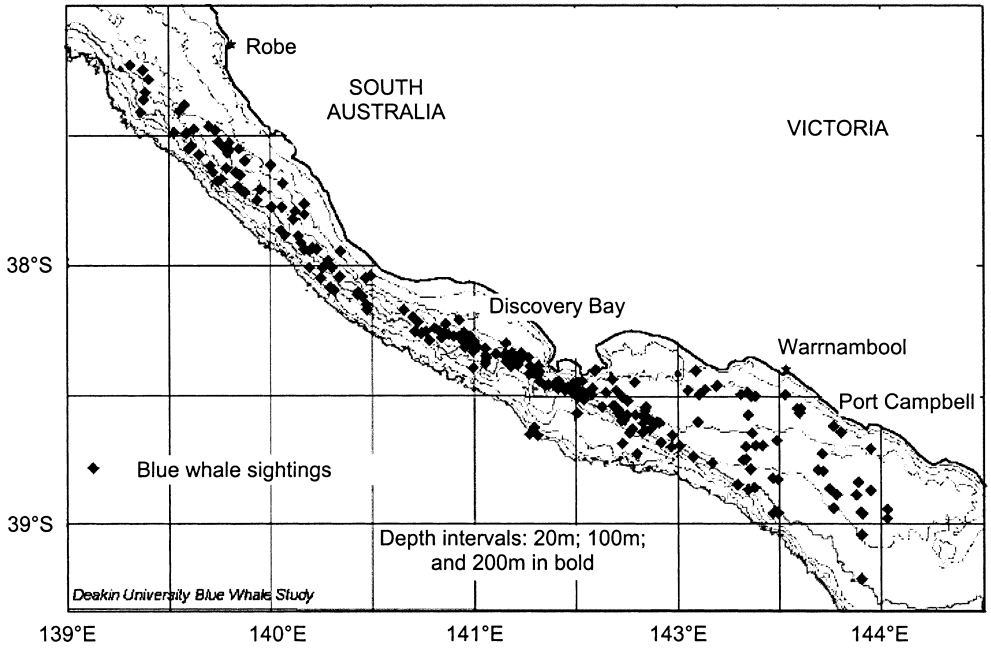


Fig. 1. Study area and blue whale sightings from the present study between December 1998 and May 2001.

Partenavia aircraft with at least two dedicated observers were flown at an altitude of 457m (1,500ft), at *ca* 140kts. Surveys were conducted in 'closing mode' i.e. whales were approached after being sighted from the trackline, to establish exact position, confirm species identity and to briefly observe behaviour. Environmental conditions, presence of other species (krill surface swarms, other marine mammals, seabird aggregations, fish schools), surface fronts, as well as fishing vessels, marine debris, oil spills and shipping were also recorded.

Boat-based studies

A yacht was used for short periods during February or March each season, for techniques including observation and videotaping of behaviour, net sampling of prey species, photo-identification, tissue biopsy/sloughed skin collection, faecal sampling, krill hydroacoustics, passive acoustic monitoring and basic oceanographic sampling.

Volunteer sighting network

Rock lobster fishermen routinely observe whales while searching for pot buoys. At the outset of the study, several were provided with a sheet showing distinguishing features of blue whales and some other whales which may occur locally (e.g. sei, humpback, southern right), and sightings record sheets. They were questioned about their sightings, including diagnostic features upon which identification was based. Sightings which could not be confirmed by the author as blue whales were discarded.

Incidental sightings

Experienced cetacean observers known to be transiting the area were asked to report blue whale observations. In addition, local charter boat operators reported some sightings. Sightings were also made from land, when blue whales were close enough to shore to be positively identified by body shape and markings.

Remote sensing

AVHRR SST (sea surface temperature) and SeaWiFs (ocean colour, or chlorophyll-*a* concentration) satellite data, and synoptic analysis weather charts, were accessed daily when available from August 1999 via the Internet (Fig. 2). These were used to monitor dynamics of prevailing winds, surface upwelling and associated primary productivity. For example, Fig. 2(a) shows that on 5 January 2000, strong SE to SSE winds were blowing along the coast. Although Figs 2(b) and 2(c) (originally colour images) lose detail and scale in the black and white rendition, they still clearly show the subsequent surface upwelling plume, as expressed in the depressed SST (7 January 2000) and enhanced chlorophyll-*a* concentrations (8 January 2000).

RESULTS

A total of 261 blue whale sightings involving 405 whales (mean pod size = 1.55, SD = 0.839; range 1-7) was recorded between 25 February 1998 and 28 April 2001.

Timing of occurrence of whales

Temporal occurrence of sightings from all aerial surveys is shown in Fig. 3(b). When compared to distance flown during aerial surveys (Fig. 3(a)), this indicates a seasonal presence/absence of blue whales, and also indicates trends in relative abundance during seasons, with more whales being sighted in March and April than in other months.

During the 1998/99 season, blue whales were first sighted on 8 December 1998 and last seen on 19 May 1999. It is possible that blue whales were present outside these times, but not recorded (this also applies to subsequent seasons). Fifty-two sightings totalling 77 blue whales (including one cow-calf pair) were made during the season. Blue whales were seen during five aerial surveys, with a maximum of 12 whales seen in any one flight. About one-third of all sightings were reported by lobster fishermen. No blue whales were sighted during aerial surveys in late May, in early August, early October or early November 1999, nor did fishermen report any sightings between May-December 1999.

During the 1999/2000 season, the first sighting was on 15 December 1999 and the last on 26 April 2000. There were 120 sightings for the season, for a total of 143 blue whales; of these, 97 sightings occurred during 13 aerial surveys. Aerial survey coverage was extended northwest and

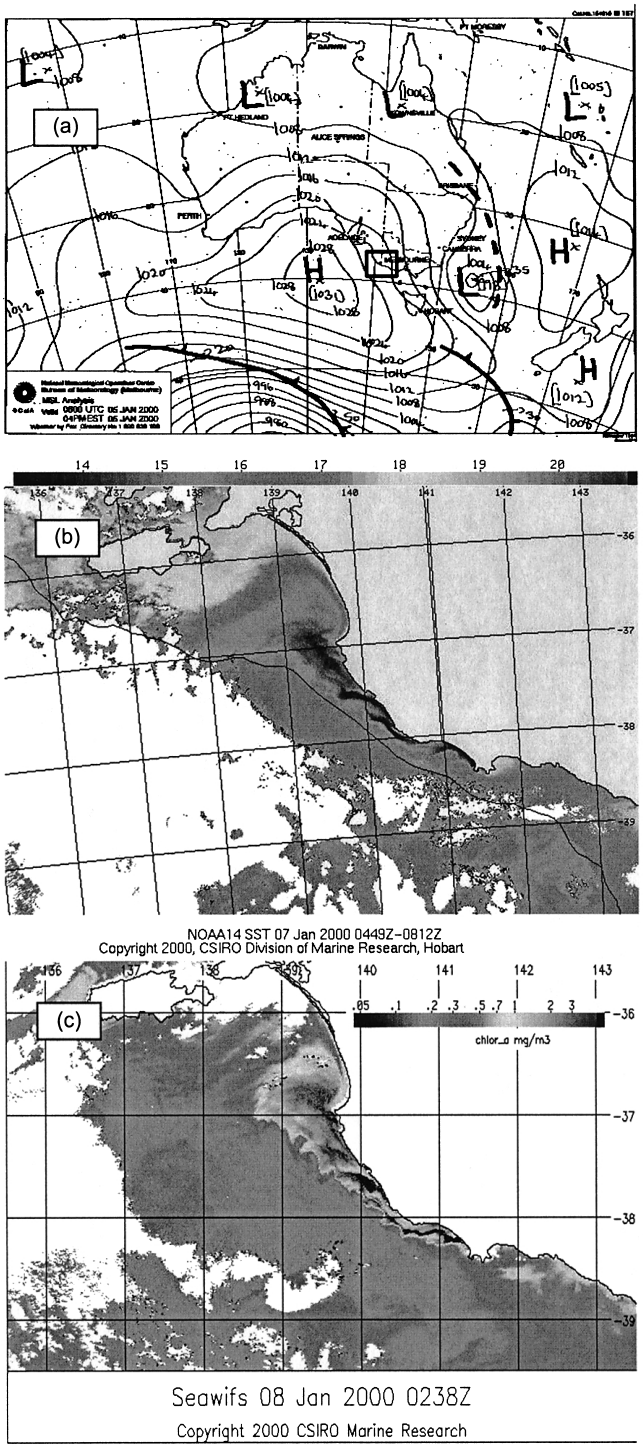


Fig. 2. (a) Typical 'upwelling' weather pattern producing strong southeast winds in the upwelling area (in box), 5 January 2000; (b) AVHRR image showing Bonney Coast upwelling, 7 January 2000; (c) SeaWiFs chlorophyll-*a* concentration, Bonney Coast upwelling, 8 January 2000. The surface upwelling plume runs northwest from the prominent cape at 141°30′E, Cape Nelson. Images courtesy of Bureau of Meteorology, CSIRO Marine Research, NASA and Orbimage.

southeast along the shelf from the previous season, roughly doubling the size of the survey area (see Fig. 3(a)). A maximum of 32 whales was seen in any one flight. Cow-calf pairs were sighted on three occasions. Only 5% of sightings were reported by fishermen, reflecting a reduced level of participation in the project by them. No blue whales were sighted during aerial surveys in August, October and November 2000, nor did fishermen report any sightings between May-December 2000.

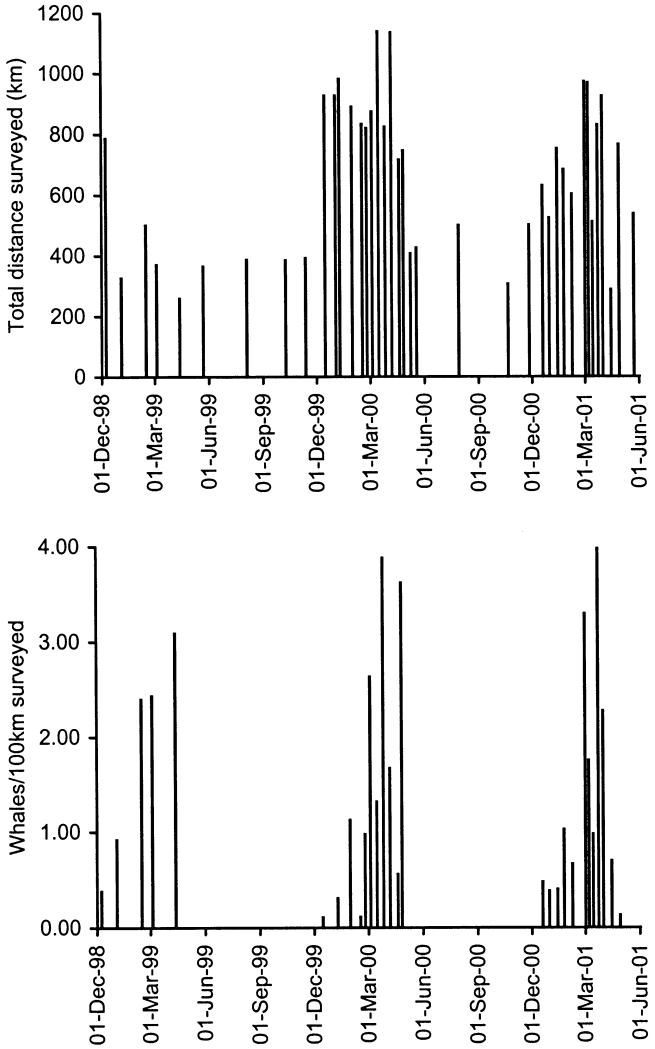


Fig. 3. (a) Timing of and total distances flown during aerial surveys; (b) Numbers of blue whales sighted per 100km flown during aerial surveys.

During the 2000/2001 season, the first sighting occurred on 19 December 2000 and the last on 28 April 2001. A total of 137 blue whales was sighted during the season, with 133 whales in 83 sightings during 13 aerial surveys. A cow-calf pair was sighted on one occasion.

The highest counts of whales occurred during some of the longest survey flights (see Figs 3(a), 3(b)). This may be explained by the fact that longer flights were possible during the period of most settled weather (March-April). This also happens to be the period when krill surface swarms are most abundant, possibly drawing greater numbers of blue whales to the area.

Distribution of sightings

Distribution of sightings is shown in Fig. 1. All sightings occurred on the continental shelf (< 200m water depth) between longitudes 139°18′E-143°03′E. Excluding this study, 45 (80%) of 56 previous blue whale sightings in Victorian waters lie within this area (Anon., 1999).

Evidence of feeding

Either feeding behaviour or defaecation were recorded in 94 (36% of total) sightings and whales were sighted within ~2km of krill surface swarms in 126 (48% of total) sightings. However on many occasions, krill swarms were sighted without blue whales in visible proximity. Obvious

feeding behaviour included energetic surface lunges to engulf prey (whales often rolled onto one side, with the mouth wide open and the throat pleats distended) and similar subsurface feeding, visible from the aircraft. More leisurely 'skim' feeding (swimming slowly through dense prey surface swarms with the mouth agape) has also been observed on several occasions (see e.g. Watkins and Schevill, 1979, for discussion of mysticete feeding methods). Whales observed with distended throat pleats and baleen visible when sighted were assumed to have just fed. During a three-hour yacht-based observation on 24 February 2000, sub-surface feeding by five blue whales, within a radius of *ca* 2km, was inferred from frequent short-duration (1-2min) fluke-up dives, with whales resurfacing near the point of diving. Hydroacoustic backscatter levels measured from a yacht were high in areas where blue whales were presumed to be feeding at depth, compared to areas surveyed where blue whales were not present. Strong backscatter signals were observed at various depths between the surface and the seafloor (90m) in an area where one blue whale was feeding (T. Pauly and P. Gill, unpublished data).

Prey

Net sampling (70cm ring net, mesh size 500μm) of surface swarms on which blue whales were directly feeding identified the whales' prey as *N. australis*. This is a neritic species reaching *ca* 20mm, and is known from New Zealand and southeast Australia (Blackburn, 1980). Surface swarms have been observed during this study in all months from October to May, and appear to be largest (some swarms exceed 1,000m in length), densest and most abundant in March and April. This species migrates vertically through the water column (Blackburn, 1980; O'Brien, 1988) and is likely to be consumed by the blue whales at greater depths (e.g. Croll *et al.*, 1998).

Upwelling and marine productivity

During late 1999, the first upwelling plume appeared in SST imagery on 17 November, following strong southeasterly winds associated with the passage of a high-pressure system (see Fig. 2 for a similar event). Further upwelling pulses followed the passage of subsequent high-pressure systems, maintaining the upwelling surface plume of cold nutrient-enriched water, which was detectable until 24 April 2000.

During late 2000, SST images showing surface upwelling first appeared on 13 November and the upwelling plume was detectable until 9 April 2001. Surface temperatures within the upwelling plume, as shown by AVHRR imagery, may be as much as 5°C cooler than surrounding waters. Chlorophyll-*a* concentrations (as shown by SeaWiFs imagery) may be elevated up to >4 mg/m^3, an order of magnitude or more above those of surrounding waters (see Fig. 2). Phytoplankton-rich upwelled waters are often visible during aerial surveys, with visible oceanic fronts clearly corresponding with those shown in satellite imagery. As shown by SeaWiFs imagery (P. Gill, unpublished data), enriched waters appear to be present for some time after active upwelling ceases, both between active upwelling pulses and after the period of upwelling-favourable wind forcing ceases.

DISCUSSION

Links between coastal upwellings, euphausiids and blue whales are apparent at a range of sites, notably in North America (Simard *et al.*, 1986; Reilly and Thayer, 1990; Sears *et al.*, 1990; Schoenherr, 1991; Croll *et al.*, 1998; 2000; Fiedler *et al.*, 1998). It appears that similar links exist in the Bonney Coast upwelling region of southern Australia. Large-area AVHRR SST images show that this seasonal upwelling zone is the most prominent in southern Australian waters (CSIRO, 2001). It occurs when high pressure cells are far enough south for their southeast wind component to be roughly parallel to this northwest-trending coast, i.e. between the months of November/December - March/April (Lewis, 1981; Schahinger, 1987). This period coincides with the presence of blue whales in the region. During late autumn, winter and spring the highs move north, prevailing winds are onshore, the upwelling is quiescent, and blue whales appear to be absent from the region.

N. australis is reported to have a production-to-biomass ratio higher than any other euphausiid (Ritz and Hosie, 1982). It is associated with an upwelling plume in New Zealand (Bradford and Chapman, 1988), and is consumed by humpback whales in New Zealand and Tasmanian waters (Dawbin, 1956; Gill *et al.*, 1998), while its congener *N. simplex* is prey of blue whales off Baja California (D. Gendron, pers. comm.). The surface distribution of *N. australis* in the present study, as determined from aerial surveys, essentially mirrors the distribution of blue whales. *N. australis* is abundant along the Bonney Coast upwelling surface plume, and the waters immediately to its east, where sub-surface upwelling is thought to occur (P. Gill, unpublished data). However, more extensive aerial surveys are needed to establish its possible distribution (and that of blue whales) outside this area. It commonly forms daytime surface swarms, at various times of year (O'Brien, 1988), as described for its congeners elsewhere (e.g. *N. simplex*, Baja California; Gendron, 1992). The surface swarming behaviour of *N. australis* has greatly facilitated this study so far by enabling the detection of swarms without the need for logistically complex hydroacoustic or net sampling surveys. However, future studies should use these methods to investigate sub-surface distribution of krill swarms. The ecology of *N. australis* in this area has yet to be properly investigated, but assuming predictability of the upwelling and its enhanced productivity during summer and autumn, it is likely that the presence and abundance of *N. australis* swarms is also seasonally predictable.

Given the direct evidence that blue whales feed regularly on *N. australis* swarms, and the large number of blue whale sightings that have been recorded since this study began, it is likely that as long as the weather patterns and oceanography which drive the seasonal upwelling in this region remain relatively stable (and human activities do not displace them), blue whales will continue to feed in this area. Fin whales (*B. physalus*; $n = 3$ sightings) and sei whales (*B. borealis*; $n = 6$ sightings), both rarely seen in Australian coastal waters, have also been observed in the feeding area. The timing of the blue whales' presence in this temperate area, which coincides with the Antarctic baleen whale feeding season, tends to support the hypothesis that these are pygmy blue whales, which are thought to mostly inhabit waters north of the Polar Front (Kato *et al.*, 1995). However, their sub-specific identity and provenance are not yet proven: so far there are no photo-identification resights with other areas, and visual, photogrammetric, genetic and acoustic techniques to discriminate between blue whale sub-species are still inconclusive (see Donovan *et al.*, 1996; IWC, 2002, pp.205-6; Ljungblad *et al.*, 1997; 1998; McCauley *et al.*, 2001).

Management issues are posed by the presence of significant oil and gas reserves in the feeding area, by a

major shipping lane intersecting it, and potentially by fisheries, and by whalewatching tourism (which is as yet undeveloped). Given the conservation status of this species, the scarcity of known feeding areas in either hemisphere and its relative accessibility (off Australia's populous south-east coast), this feeding area offers a rare opportunity to increase knowledge of blue whale biology and ecology and that of their prey. It also provides management agencies with an opportunity to devise and implement management initiatives which will contribute towards the continued survival of blue whales in this region.

ACKNOWLEDGEMENTS

Sincere thanks are due to all those who have assisted the study in various ways. Tim Edkins, Peter Drew, Bill Doble, Peter Manthorpe and other crew of *Iniquity* provided invaluable assistance with yacht-based fieldwork. Deborah Thiele, Shannon McKay, Graeme McKenzie, Margie Morrice and Giorgio Palmeri were particularly helpful with aerial surveys. Portland lobster fishermen Dave Johnston, Steve Nathan, Doc Andre, Neil O'Connell and Chad Hall contributed valuable sightings data. Deborah Thiele, Brad Mitchell, Peter Fairweather, Andrew Levings and Margie Morrice provided advice and encouragement. George Cresswell of CSIRO provided insights into the Bonney Coast upwelling. Remote sensing data was generously made available by CSIRO Marine Research, NASA and Orbimage.

The study was funded by Environment Australia (Natural Heritage Trust); the W.V. Scott Trust; Deakin University; the Australian Geographic Society' and Woodside Energy Pty Ltd. Biopsy in 2000 was conducted under Whale Protection Act Scientific Permit P1999/07, issued by Environment Australia. This paper was improved by comments from Nick Gales and Steve Nicol of the Australian Antarctic Division, and from two anonymous reviewers. This is a contribution from the Deakin University Marine and Migratory Species Research Group, and forms part of work towards a PhD at Deakin University, supervised by Peter Fairweather and Brad Mitchell.

REFERENCES

Anon. 1999. *Atlas of Victorian Wildlife Database*. Arthur Rylah Institute, Parks, Flora and Fauna Division, Department of Natural Resources and Environment, Melbourne.

Bannister, J.L., Kemper, C.M. and Warneke, R.M. 1996. *Wildlife Australia, Endangered Species Program.* The Action Plan for Australian Cetaceans. Project Number 380. Australian Nature Conservation Agency, Canberra. 242pp. [Available from: *www.environment.gov.au/library/pubs*].

Blackburn, M. 1980. Observations on the distribution of *Nyctiphanes australis* Sars (Crustacea, Eupausiidae) in Australian waters. CSIRO Australian Division of Fisheries and Oceanography, report no. 119. (Unpublished). 10pp. [Available from: CSIRO Marine Research, Hobart, Tasmania].

Bradford, J.M. and Chapman, B. 1988. *Nyctiphanes australis* (Euphausiacea) and an upwelling plume in western Cook Strait, New Zealand. *NZ J. Mar. Freshwater Res.* 22:237-47.

Branch, T.A. and Butterworth, D.S. 2001. Estimates of abundance south of 60^0S for cetacean species sighted frequently on the 1978/79 to 1997/98 IWC/IDCR-SOWER sighting surveys. *J. Cetacean Res. Manage.* 3(3):251-70.

Clapham, P.J., Young, S.B. and Brownell, R.L., Jr. 1999. Baleen whales: conservation issues and the status of the most endangered populations. *Mammal Rev.* 29:35-60.

Commonwealth of Australia. 1999. Environment Protection and Biodiversity Conservation Act 1999. Act of Parliament No. 91, 1999. Parliament House, Canberra.

Croll, D.A., Tershy, B.R., Hewitt, R.P., Demer, D.A., Fiedler, P.C., Smith, S.E., Armstrong, W., Popp, J.M., Kiekhefer, T., Lopez, V.R., Urban, J. and Gendron, D. 1998. An integrated approach to the foraging ecology of marine birds and mammals. *Deep-Sea Res. II* 45:1353-71.

Croll, D.A., Marinovic, B., Benson, S., Chavez, F.P., Black, N., Termullo, R. and Tershy, B.R. 2000. From wind to whales: trophic links in an upwelling ecosystem. Final report to the Monterey Bay National Marine Sanctuary, Contract no. 50ABNF500153. (Unpublished). 46pp. [Available from: D.A. Croll, Institute of Marine Science, University of California at Santa Cruz, Santa Cruz, California].

CSIRO. 2001. Interactive composite (SST) image of Australia. Remote sensing web-site (*http://www.marine.csiro.au/~lband/webpoint*), CSIRO Marine Research, Hobart.

Dawbin, W.H. 1956. The migration of humpback whales which pass New Zealand. *Trans. R. Soc. NZ* 84(1):147-96.

Donovan, G., Hammond, P., Larsen, F., Reilly, S., Kato, H., Clark, C. and Brownell, R.L. 1996. Report of the Scientific Committee, Annex E. Report of the sub-committee on Southern Hemisphere baleen whales, Appendix 2. An approach for responding to the Commission's resolution. *Rep. int. Whal. Commn* 46:126-7.

Fiedler, P.C., Reilly, S.B., Hewitt, R.P., Demer, D., Philbrick, V.A., Smith, S., Armstrong, W., Croll, D.A., Tershy, B.R. and Mate, B.R. 1998. Blue whale habitat and prey in the California Channel Islands. *Deep-Sea Res. II* 45:1781-801.

Gendron, D. 1992. Population structure of daytime surface swarms of *Nyctiphanes simplex* (Crustacea: Euphausiacea) in the Gulf of California, Mexico. *Mar. Ecol. Prog. Ser.* 87:1-6.

Gill, P.C., Evans, K.J. and Wapstra, H. 1998. Feeding by humpback whales in Tasmanian waters. *Rec. Q. Vic. Mus. (Launceston)* 107:1-5.

Herzfeld, M. 1997. The annual cycle of sea surface temperature in the Great Australian Bight. *Prog. Oceanogr.* 39:1-27.

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*. University of California Press, Berkeley and Los Angeles. xv+789pp.

International Whaling Commission. 2002. Report of the Scientific Committee. Annex G. Report of the Sub-Committee on the Comprehensive Assessment of Whale Stocks - In-Depth Assessments. *J. Cetacean Res. Manage. (Suppl.)* 4:192-229.

Kato, H., Miyashita, T. and Shimada, H. 1995. Segregation of the two sub-species of the blue whale in the Southern Hemisphere. *Rep. int. Whal. Commn* 45:273-83.

Kato, H., Bannister, J., Burton, C., Ljungblad, D., Matsuoka, K. and Shimada, H. 1996. Report on the Japan/IWC blue whale cruise 1995-96 off the southern coast of Australia. Paper SC/48/SH9 presented to IWC Scientific Committee, June 1996, Aberdeen (unpublished). 35pp. [Paper available from the Office of this Journal].

Lewis, R.K. 1981. Seasonal upwelling along the south-eastern coastline of South Australia. *Aust. J. Mar. Freshwater Res.* 32:843-54.

Ljungblad, D.K., Stafford, K.M., Shimada, H. and Matsuoka, K. 1997. Sounds attributed to blue whales recorded off the southwest coast of Australia in December 1995. *Rep. int. Whal. Commn* 47:435-9.

Ljungblad, D.K., Clark, C.W. and Shimada, H. 1998. A comparison of sounds attributed to pygmy blue whales (*Balaenoptera musculus brevicauda*) recorded south of the Madagascar Plateau and those attributed to 'true' blue whales (*Balaenoptera musculus*) recorded of Antarctica. *Rep. int. Whal. Commn* 48:439-42.

McCauley, R.D., Jenner, C., Bannister, J.L., Burton, C.L.K., Cato, D.H. and Duncan, A. 2001. Blue whale calling in the Rottnest Trench - 2000, western Australia. *J. Acoust. Soc. Am*: Submitted.

O'Brien, D.P. 1988. Surface schooling behaviour of the coastal krill *Nyctiphanes australis* (Crustacea: Euphausiacea) off Tasmania, Australia. *Mar. Ecol. Prog. Ser.* 42:219-33.

Palacios, D.M. 1999. Blue whale (*Balaenoptera musculus*) occurence off the Galapagos Islands, 1978-1995. *J. Cetacean Res. Manage.* 1(1):41-51.

Reilly, S.B. and Thayer, V.G. 1990. Blue whale (*Balaenoptera musculus*) distribution in the eastern tropical Pacific. *Mar. Mammal Sci.* 6(4):265-77.

Ritz, D.A. and Hosie, G.W. 1982. Production of the euphausiid *Nyctiphanes australis* in Storm Bay, south-eastern Tasmania. *Mar. Biol.* 68:103-8.

Schahinger, R.B. 1987. Structure of coastal upwelling events observed off the south-east coast of South Australia during February 1983-April 1984. *Aust. J. Mar. Freshwater Res.* 38:439-59.

Schoenherr, J.R. 1991. Blue whales feeding on high concentrations of euphausiids around Monterey Submarine Canyon. *Can. J. Zool.* 69:583-94.

Sears, R., Williamson, J.M., Wenzel, F.W., Bérubé, M., Gendron, D. and Jones, P. 1990. Photographic identification of the blue whale (*Balaenoptera musculus*) in the Gulf of St. Lawrence, Canada. *Rep. int. Whal. Commn* (special issue) 12:335-42.

Simard, Y., de Ladurantaye, R. and Therriault, J.-C. 1986. Aggregation of euphausiids along a coastal shelf in an upwelling environment. *Mar. Ecol. Prog. Ser.* 32:203-15.

Watkins, W.A. and Schevill, W.E. 1979. Aerial observations of feeding behavior in four baleen whales, *Eubalaena glacialis, Balaenoptera borealis, Megaptera novaeangliae* and *Balaenoptera physalus*. *J. Mammal.* 60(1):155-63.

Errors in age estimates of North Atlantic minke whales when counting growth zones in *bulla tympanica*

ERIK OLSEN

Division of Resource Ecology, Institute of Marine Research, PO Box 1870 Nordnes, N-5817 Bergen, Norway
Contact e-mail:eriko@imr.no

ABSTRACT

Age estimation of common minke whales (*Balaenoptera acutorostrata*) has always been difficult, and the accuracy of the current method of counting growth layer groups (GLGs) in the periosteal zone of the *tympanic bulla* has been questioned. To evaluate this method, two readers aged three sections from each of right and left *bulla tympanica* from 35 male and 57 female North Atlantic minke whales. A single age estimate with variance was calculated for each whale using General Linear Mixed Model Poisson based regression, and this estimate was compared with the number of ovulations and body length to evaluate the bias of the age determination method. The results showed a poor fit between age and number of ovulations with $R^2 = 0.0014$. Bias was estimated to be a 37% underestimate of 'true' age assuming an ovulation rate of one per year and age at sexual maturity of eight years. Precision of the bulla age estimates was lower than those of Antarctic minke whales aged using the earplug method. The high bias reduces the applicability of the bulla method in routine age-determination with a management objective. Other age determination methods for the species should be improved or developed to ensure proper monitoring of demography and life history for the North Atlantic minke whale.

KEYWORDS: COMMON MINKE WHALE; ATLANTIC OCEAN; AGE DETERMINATION; AGE AT SEXUAL MATURITY

INTRODUCTION

Precise, unbiased age estimates are important requirements in managing different stocks of terrestrial and aquatic organisms (Morris, 1972; Birney *et al.*, 1975). Age is especially important when one wishes to study the age structure, life-history and catch-at-age history of the population in question (Nuckle and Bergeron, 1983). Determining the correct age of an individual is often difficult and errors can lead to an incorrectly assigned age (Bowering and Nedraas, 2001; Richards *et al.*, 1992; Olsen and Skaug, 2002). The difference between the true age *a* and the estimated age *A* is a measure of the bias of the age estimate, while the variance of several independent age estimates of the same individual is a measure of the precision of the estimate. Poor precision and high bias can each lead to errors in the demographic parameters required for study (Schnute and Richards, 1995), and it is important that these are investigated and quantified. The North Atlantic common minke whale stock is currently exploited by Norwegian whalers and Greenland Inuit hunters (IWC, 1999), and estimating correct age has been a problem for a long time. In the Norwegian minke whale studies, less than 20% of the animals could be aged using ear plugs (Christensen, 1992) mainly because of difficulties with plug removal and collection. Instead, whales from the Norwegian catch have been aged by counting the annual growth layer groups (GLGs) in the periosteal layer of the ear bone, *bulla tympanica* (Christensen, 1981). Two early follow-up analyses indicated that there was an error in ageing when using the bulla method (Larsen and Kapel, 1983; Sukhovskaya *et al.*, 1985), but the error was not quantified until Christensen (1995) found an 80% agreement (± 1 year) between three readers. However, a later study by Olsen (1997) indicated that the bulla method underestimated the 'true' age. These disagreements necessitated the study of errors in age estimates based on the bulla method with the aim of quantifying both precision and bias of the method. A project was started in 1999 with three aims: (1) evaluate the bias of the method through a multi-reader experiment; (2) estimate the precision of bulla ageing method through a multi-reader experiment; (3) evaluate the reliability of the bulla method when used in routine age estimation. The focus of the present study is on bias (1) and use of the bulla method in routine age estimation (3). The question of precision (2) has been dealt with in detail in Olsen and Skaug (2002).

Bias

The bias of a single age estimate is the difference between estimated age *A* and true age *a*, and the ideal method of studying this is by conducting a study on known-age animals. For practical reasons this is impossible for minke whales, and one is left with two alternatives: (1) to compare bulla age estimates with other independent estimates of age (i.e. age estimates based on alternative methods); or (2) to compare bulla age estimates with independent growth parameters. No validated alternative ageing method currently exists for North Atlantic minke whales, which leaves comparing bulla age with different measures of relative age i.e. number of ovarian scars (ovulations) and length.

Fin whales (*Balaenoptera physalus*) and Antarctic minke whales (*Balaenoptera bonaerensis*) have been shown to have a regular ovulation rate throughout life (Laws, 158; Ohsumi and Masaki, 1975), and from studies of pregnancy rates (Christensen, 1975; Olsen, 1997) North Atlantic minke whales seem to share this characteristic. After ovulation, the ruptured follicle increases in size and becomes a *corpus luteum*, which lasts through pregnancy and lactation, after which it shrinks to a smaller size (minimum size is less than 1cm in North Atlantic minke whales) and is called a *corpus albicans* which persists for life (Laws, 1958). Thus, the ovaries of sexually mature minke whales have a permanent record of the ovulation history of the individual. Comparing the true age of minke whales with the number of ovulations (found by counting *corpus lutea* and *alibicans* in the ovaries) would yield a plot fitting within the area delimited by the error bars in Fig. 1. The error bars delimit the range of number of ovulations for a whale at a given age based on the ovulation rate. The upper limit is based on an ovulation rate of two ovulations per year, while the lower is based on a biennial ovulation rate, which are realistic upper and lower limits for the species. The 'X's plotted in the middle are based on an annual ovulation rate, which is the best current

estimate for North Atlantic minke whales based on the high (80-98%) annual pregnancy rates of mature females (Jonsgård, 1951; Christensen, 1975; Olsen, 1997). However, a higher ovulation rate is not unlikely considering that in general, mammals can have spontaneous abortions in early pregnancy and new ovulation usually follows (Laws, 1958).

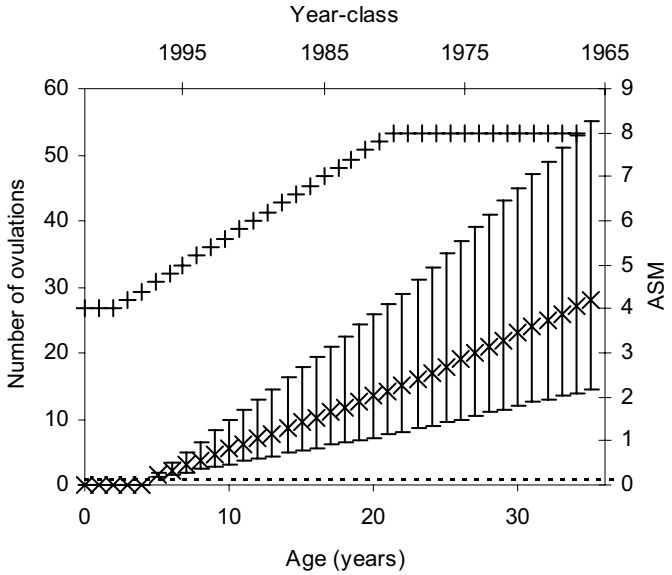


Fig. 1. Theoretical increase in the numbers of ovulations with age of female minke whales given an annual ovulation rate, and a reduction in age at sexual maturity (ASM) from eight years for the 1980 year-class to four years for the 2000 year-class. The error bars represent limits given an ovulation rate of $2*year^{-1}$ (upper bar) and $0.5*year^{-1}$ (lower bar). The first ovulation is indicated by the hatched horizontal line. X = ovulations; + = ASM.

To model a possible density-dependent decline in age at attainment of sexual maturity (ASM) the number of ovulations plotted in Fig. 1 are corrected for a decrease in ASM from eight years in 1980 to four years in 2000, where knife-edge maturation for all year-classes is assumed. The upper and lower limits of ASM were chosen deliberately and deemed realistic as ASM was estimated to eight years (±0.7 years, 95% confidence interval) by Olsen (1997), and recently to 5.8 years (±3 years, 95% confidence interval) by Olsen and Sunde (2002). Fig. 1 shows that regardless of variable ovulation rate or a decline in ASM an increase in ovulations with age could be expected if the age estimates were accurate (unbiased).

Body length increases with age, and can be a useful indicator of relative age when monitoring young year-classes as these can be identified and annual growth measured. However, mammals do not grow forever, and generally reach a maximum body size soon after attainment of sexual maturity (e.g. Brody, 1945). The sexes also have different growth patterns and different maximum body lengths. In this case female minke whales grow markedly larger than the males. In addition, there are large individual differences in physical growth. Physical growth can be modelled by using one of several non-linear growth equations. One of these is the von Bertalanffy equation (Equation 1), which was used with success by George *et al.* (1999) and Olsen and Sunde (2002) to model body length versus age estimated using the aspartic acid racemisation technique (Olsen and Sunde, 2002). To facilitate comparison with these and other studies, this growth model was used in the analysis.

Von Bertalanffy growth equation:

$$Length = L_{MAX} - L_{MAX} \times e^{(k \times t_0 - k \times age)} \quad (1)$$

L_{MAX} is the maximum body length, k is the growth rate, and t_0 is the age at length 0.

Fig. 2 plots minke whale body length with age based on published estimates of length at birth (280cm, as presented by Jonsgård (1951)) and sex-specific maximum body length (male = 812, female = 840), calculated from the 20% largest males and females in the data (518 male and 1,264 female) used by Olsen (1997).

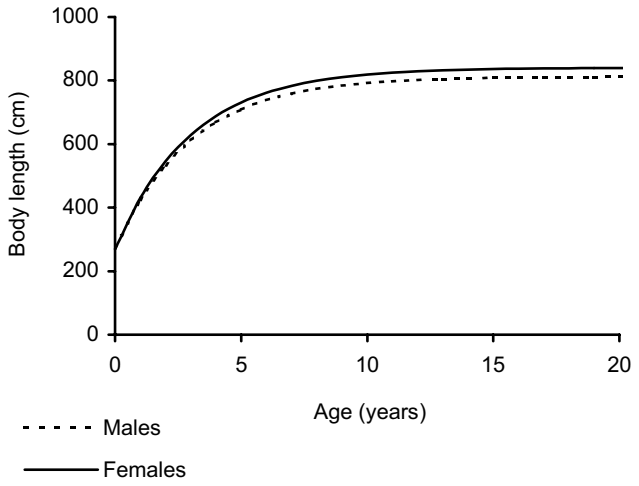


Fig. 2. Physical growth of male and female minke whales modelled by the von Bertalanffy growth equation where L_{MAX} was calculated from data in Olsen (1997). The parameter k (growth rate) was the same as estimated by (Olsen and Sunde, In press), while t_0 was chosen such that length at birth ≈ 280cm (from Jonsgård, 1951).

Precision

Precision, or random error, is a measure of how close agreement there is between parallel age estimates of the same whale. Precision is dependent on biological factors determining the formation of GLGs in bullae, and on the subjective choices made by the reader when analysing a particular bulla. It is of interest to quantify the error as a whole, and what part is caused by biological factors and what is caused by reader variability. Olsen and Skaug (2002) developed a method to estimate the variance of multiple age estimates and divided this into different components, to then make an estimate of the age of an individual. The same method is employed in the present study to estimate the age and variability based on multiple readings of bulla GLGs.

MATERIALS AND METHODS

Right and left bullae were collected from 35 male and 57 female minke whales caught onboard two Norwegian whaling vessels operating in the North Sea and Central Atlantic (ES and CA, IWC small management areas, see IWC, 1994) in May-July 1999 and 2000 (Table 1). From 28 of these, only one bulla was collected as the other was destroyed after the grenade on the harpoon detonated, or it was inaccessible during flensing. The length distribution of the sampled whales is shown in Fig. 3. At the laboratory the bullae were cleaned and 0.2mm thick transverse segments were cut from the medial part of each bulla using a dual-bladed circular saw. Three such segments were cut

Table 1

Whale number, minimum and maximum age read, number of parallel age readings from a whale (*n*), sex, number of ovulations (ov.), standard body length, estimated age (based on GLMM regression) and standard deviation of the age estimate from reading left and right *bulla* from 92 minke whales caught in the North Sea in 1999 and 2000. An * after 'whale' indicates that only one *bulla* was collected.

North Sea								
Whale	Min.	Max.	*n*	Sex	Ov.	L(cm)	Est. age	SD Est.
1999								
F10	2	20	12	F	13	825	9.1	0.90
F11	3	17	12	F	0	790	7.7	0.83
F12	5	33	12	F	12	800	13.9	1.11
F14	8	18	12	M	0	835	11.5	1.01
F15	4	19	12	M	0	845	10.0	0.94
F16	5	18	12	M	0	811	10.6	0.97
F17*	3	15	6	F	0	666	9.6	1.31
F18*	3	19	6	F	0	660	13.7	1.56
F19*	3	11	6	F	0	697	6.8	1.10
F2*	3	15	6	F	1	760	8.1	1.21
F20*	8	18	6	F	0	862	12.1	1.47
F21*	6	18	6	M	0	785	10.9	1.40
F22*	5	12	6	F	0	840	7.8	1.18
F23*	3	12	6	M	0	733	7.8	1.18
F24*	3	12	6	F	7	855	6.3	1.06
F25	9	16	12	F	1	731	12.8	1.07
F26	9	24	10	F	16	860	14.1	1.21
F27*	3	14	5	M	0	728	6.8	1.18
F28	5	11	10	M	0	810	6.3	0.81
F29	8	22	10	F	0	535	11.8	1.10
F3*	7	17	6	F	8	840	9.8	1.32
F30	2	9	10	M	0	745	5.2	0.73
F31	4	12	10	M	0	745	6.4	0.81
F4	6	22	12	F	5	820	12.7	1.07
F5	4	32	12	F	4	800	8.5	0.87
F6	3	22	12	M	0	810	7.3	0.81
F7	2	11	12	F	0	760	5.7	0.71
F8	3	15	12	F	5	745	9.0	0.89
F9	6	24	12	F	3	780	11.6	1.02
K18	4	11	12	M	0	769	6.8	0.78
K19	3	16	12	M	0	762	7.1	0.79
K20	4	20	12	M	0	730	8.7	0.88
K21	6	19	12	F	0	675	9.7	0.93
K22	6	25	12	F	0	725	12.8	1.07
K23	8	21	12	F	0	726	13.7	1.11
K24	0	2	12	M	0	485	0.5	0.22
K25	7	29	12	M	0	840	14.0	1.12
K26	7	17	12	M	0	740	10.8	0.98
K27	9	26	12	F	6	760	13.1	1.08
K28	5	32	12	F	10	835	12.2	1.04
K29	8	19	12	M	0	832	11.4	1.01
K30	7	14	12	F	4	822	9.9	0.94
K31	5	20	12	M	0	820	10.8	0.98
K33	4	14	12	M	0	700	7.8	0.84
K34	4	14	12	F	6	788	9.0	0.89
K35	4	22	12	M	0	845	9.6	0.93
2000								
K10	6	18	12	F	12	835	10.2	0.95
K11	4	13	12	M	0	690	8.3	0.86
K12	8	22	12	M	0	810	12.1	1.04
K28	3	13	12	M	0	810	9.1	1.14
K29	0	5	12	F	0	540	1.7	0.90
K30*	7	11	6	M	0	815	8.5	1.23
K31*	6	10	6	F	0	880	7.8	1.18
K9	5	15	12	M	0	780	8.1	0.85

Central Atlantic								
Whale	Min.	Max.	*n*	Sex	Ov.	L(cm)	Est. age	SD Est.
1999								
K1	5	23	11	F	12	857	8.5	0.87
K10	6	13	12	F	5	845	10.0	0.95
K11*	7	13	6	F	0	705	10.4	1.36
K12	2	12	12	F	7	802	6.7	0.77
K13	4	12	12	F	18	855	7.5	0.82
K14	5	17	12	M	0	730	10.4	0.96
K15	9	19	6	F	4	858	11.3	1.42
K17	5	18	12	F	0	760	10.2	0.95
K2	7	16	12	F	9	840	10.2	0.95
K3	7	27	12	F	11	810	12.2	1.04
K4	9	14	12	F	15	885	11.1	0.99
K5	6	21	12	F	10	870	10.2	0.95
K6	8	13	12	F	3	776	9.4	0.91
K7	5	11	12	F	15	855	8.1	0.85
K8	4	11	12	F	10	801	7.2	0.80
K9	0	14	12	F	11	868	8.1	0.85
2000								
K1	4	16	12	M	0	747	8.8	0.89
K13*	4	10	6	M	0	585	7.3	1.14
K14*	3	13	6	F	0	700	8.0	1.19
K15*	7	14	6	F	0	765	8.3	1.22
K16*	9	22	6	F	0	765	12.2	1.48
K17*	13	21	6	F	0	770	16.7	1.72
K18*	10	13	6	F	0	848	11.6	1.44
K19*	7	11	6	F	0	670	9.0	1.26
K2	6	16	12	M	0	770	10.6	0.97
K20*	7	23	6	F	0	740	13.9	1.57
K22*	3	10	6	M	0	792	6.2	1.05
K23*	4	13	6	F	0	770	8.5	1.23
K24*	12	21	6	F	0	715	15.2	1.65
K25*	7	12	6	F	0	810	9.1	1.28
K26*	6	20	6	F	0	805	12.9	1.52
K27*	5	11	6	M	0	720	7.3	0.39
K3	5	13	12	M	0	712	7.2	0.80
K4	5	18	12	M	0	793	9.9	0.94
K5	5	17	12	F	1	725	10.9	0.98
K6	6	14	12	F	12	820	10.3	0.96
K7	6	13	12	M	0	735	8.8	0.89
K8	5	12	12	M	0	820	6.7	0.77

from each bulla, all within 1cm of each other. The three segments were mounted on one microscope-slide. All slides were randomly given new numbers to prevent the readers from recognising the slides between readings. Two readers aged each bulla segment independently, where each reader first read the first section of all bulla. The slides were then put away for 5-7 days, the slides were renumbered, and the readers proceeded to read the second and then the third segment with renumbering and rest in-between.

Previous studies have not indicated any morphological difference between the North Sea and Central Atlantic (Christensen *et al.*, 1990) and the data were therefore pooled from both areas. Using a modified Poisson model and GLMM regression (Olsen and Skaug, 2002) a single age estimate was made for each whale based on all readings from the particular animal. The model was also used to estimate the standard deviation of this estimate. The model estimated ages were compared with the body length and ovulations.

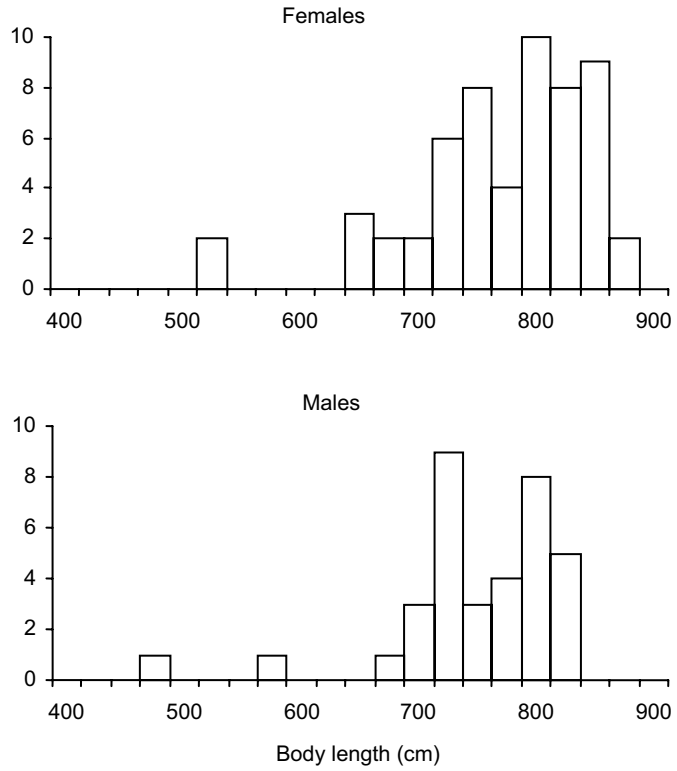


Fig. 3. Histograms of body length (cm) of male and female minke whales sampled in the North Sea and Central Atlantic in 1999 and 2000.

When analysing age versus ovulations a linear regression line was fitted to the plot to evaluate if there was a correlation and if there was a linear increase of one ovulation per year of age as might be expected. For the age versus length analysis, the sample was split into male and female sub-samples (since each sex shows a different growth pattern), a von Bertalanffy growth equation was fitted to the plot and $L_{MAX,}$ k, and t_0 were estimated.

RESULTS

The length distribution of the sample is shown in Fig. 3, and as expected shows that females achieve longer maximum body lengths than the males. Samples from the smaller length groups (< 650cm) were few, with only five females and two males. This is not representative of the population, but an artefact due to the size-selective catching by the whalers who have a set catch limit and wish to maximise the amount of meat. The sex ratio for the whole sample set was 38% males and 62% females. However splitting the sample by region revealed that in the North Sea, the sex ratio was 46% males and 54% females, whilst in the Central Atlantic it was 26% males and 74% females. The different sex ratios are probably not an artefact of the hunting, but rather a result of the sex-segregated migration of the species (Jonsgård, 1951).

Even with 28 whales missing one bulla completely, all whales in the sample were given at least six independent age estimates (three by each reader). For six whales, one of the readers was unsure about the estimate of a particular bulla segment and therefore did not age that particular segment.

Age estimates of single bulla segments ranged from 0-33 years respectively (Table 1), and the largest difference between minimum and maximum age estimates for an individual whale was 28 years (whale F5/1999). The standard deviation of the modelled age estimates ranged from 0.2-1.7, translating into 95% confidence intervals of the age estimates from ± (0.1-1.4) years. Plotting the coefficient of variation (CV) versus age (Fig. 4) showed a slight decrease in CV with age, except for the two youngest animals who had a markedly higher CV than the rest of the sample. As expected the whales with a missing bulla had a higher CV than the others.

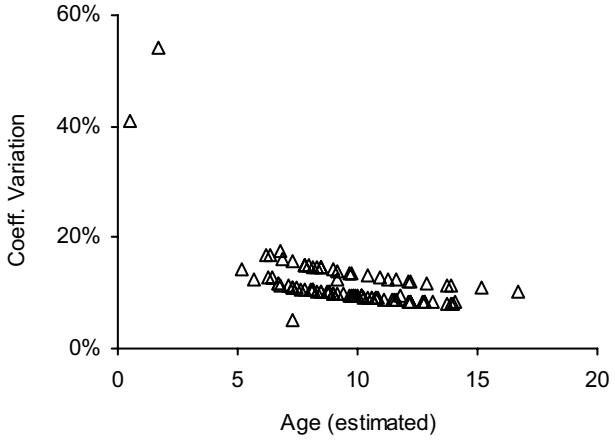


Fig. 4. Estimated age vs coefficient of variation of bulla age estimates from 57 female and 35 male minke whales. For 28 whales either the left or right bulla was missing, and this reduced the sample size for these whales by half, indicated in the plot by the points with higher CV than whales with similar age.

In the analysis of bias, no correlation was found between the estimated age and number of ovulations (Fig. 5, $R^2 = 0.001$) because the number of ovulations at a given estimated age ranged from 0 to ~15. This was contrary to expectations of a linear increase in number of ovulations after sexual maturity (see Fig. 1) and can only be explained by the bulla age underestimating true age. The underestimation was also apparent from the plot of residuals shown in Fig. 6 indicating a mean underestimation of 37% given ASM of eight years and 12% given ASM of four years.

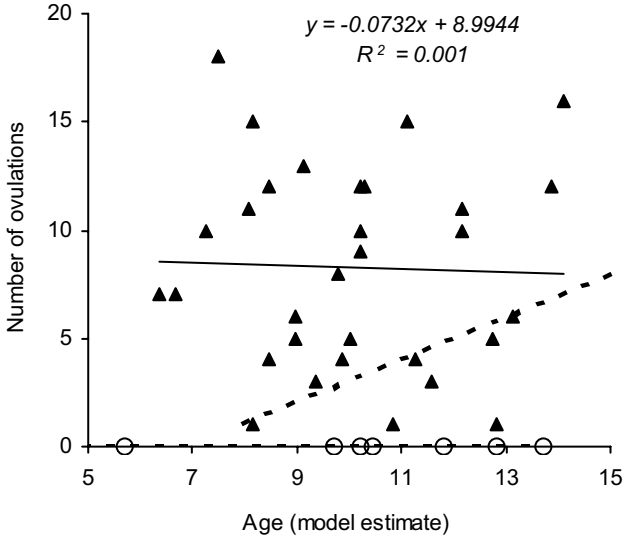


Fig. 5. Observed relationship between age and number of ovulations from 92 minke whales sampled in the North Sea and Central Atlantic in 1999 and 2000. A linear regression line is fitted to the whales with one or more ovulations (filled triangles) with regression equation and R^2 value is shown in the plot. The dashed line indicates the theoretical expected relationship given an annual ovulation rate and knifes-edge maturation at eight years. One immature whale estimate to one year of age is not shown.

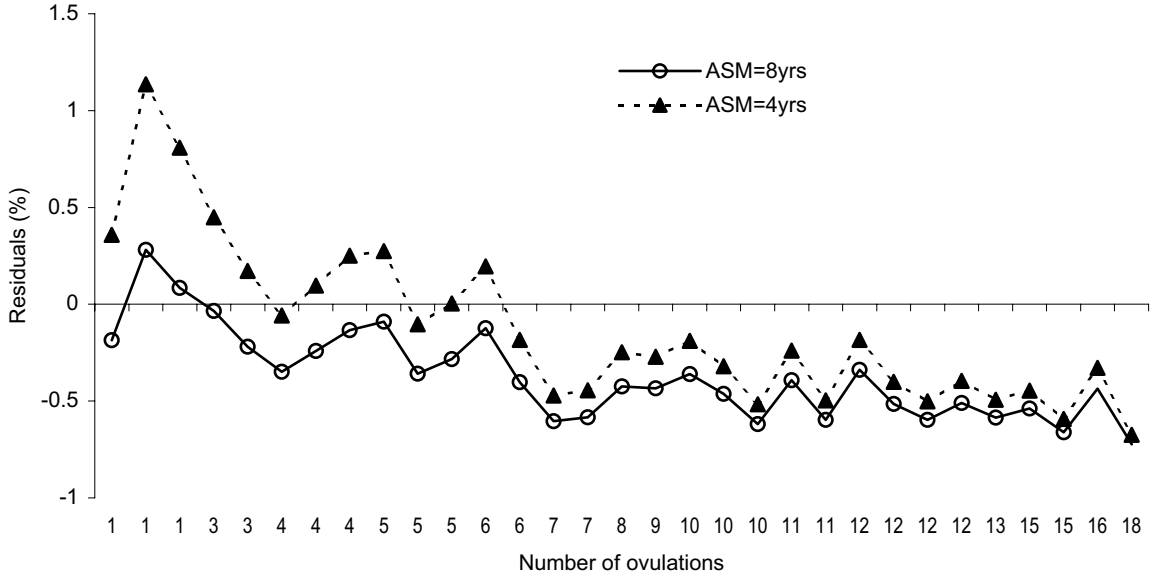


Fig. 6. Mean percentage residuals of the age estimates of individual whales when compared with the theoretical age estimated given annual ovulation rate and knifes-edge maturation at eight years (circles), or four years (triangles).

Using either assumed ASM indicated a progressive underestimation of age, with the youngest age classes being overestimated while the older being underestimates.

When plotting body length versus age (Fig. 7 and Fig. 8) it was found that the male sample seemed to fit the von Bertalanffy growth equation better than females. From Fig. 8 it can be seen that one whale is apparently an outlier (F29/1999), but removing this from the analysis did not notably improve or change the fit of the von Bertalanffy growth model. The curves were widely different for both sexes, which was apparent both from visual inspection and from comparing the model parameters (Table 2). These revealed that the males would grow to a larger size than females, but this was a result of different fit due to the lack of small animals and varying ages estimated for a give length class. Most whales in the sample were full-grown, or close to it, indicating little increase in size with age. Thus, the von Bertalanffy curves were driven by the few small and young animals in the samples introducing large differences between the sexes in model fit. In fact, the dataset encompassed only four animals less than 600cm long, and only one male and female was estimated to be less than five years of age.

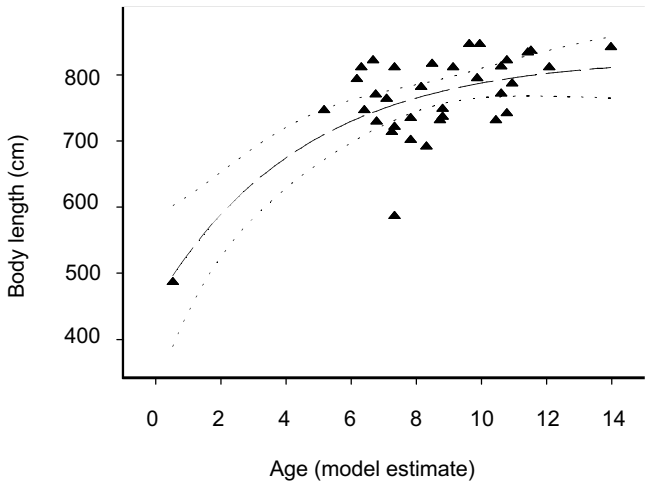


Fig. 7. Body length vs estimated age for 35 male, North Atlantic minke whales caught in 1999 and 2000. A von Bertalanffy growth equation with 95% confidence intervals is fitted to the data.

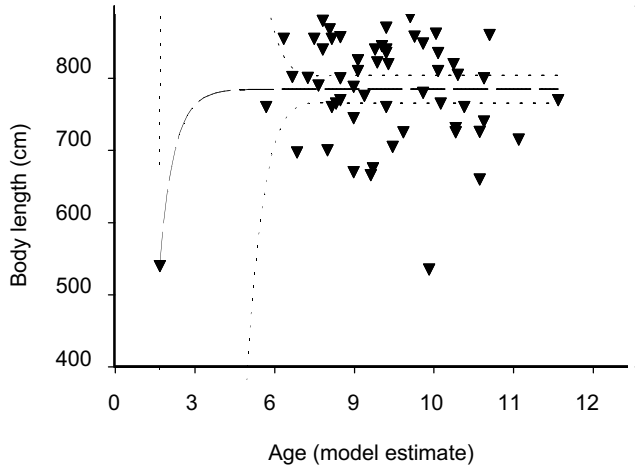


Fig. 8. Body length vs estimated age for 57 female, North Atlantic minke whales caught in 1999 and 2000. A von Bertalanffy growth equation with 95% confidence intervals is fitted to the data.

Table 2

Parameter estimates for the fitted von Bertalanffy growth equations in Figs 6 and 8. L_{MAX} is an estimate of mean maximum body length, k governs the growth rate while t_0 is a theoretical estimate of age at 0 length.

Sex	L_{MAX}	*k*	t_0
Males	827	0.22	-3.6
Females	785	1.81	1.0

DISCUSSION

An age determination method can only be useful in management if it yields fairly precise and unbiased estimates of true age without the need for resorting to accessory information to estimate age. In addition, the method must be applicable for both sexes and to most of the age-spectrum in the population. This investigation was devised with these considerations in mind, with all bulla age readings being blind readings, and attempting to use as representative a dataset as was possible to obtain. Few small whales (< 600cm body length) were included and any further study

should attempt to sample more small whales. However, the sampled whales included females who had had 0-18 ovulations, corresponding to ages < 7-25 years given an ASM of ~8 years (Olsen, 1997).

The precision estimated for the GLMM model age estimates ranged from ± (0.1-1.4) years (95% confidence interval), which was consistent with Christensen (1995) who found 80% agreement within ±1 year for three readers. Coefficients of variations decreased with increasing age (Fig. 4), and ranged from 5-14% for whales with ages > 5 years based on 12 parallel readings. The sample included two animals with an estimated age of < 2 years (K24/1999 and K29/2000). The ages of the other whales were estimated to be > 5 years, making the two smallest outliers in the GLMM regression analysis, resulting in very high CVs for these two whales. The GLMM model suffered from a non-homogenous dataset, lacking equal dispersal of age estimates of the sampled whales. Even so, the mean CVs in the present study were larger than those found by (Kato *et al.*, 1991b) from two parallel readings of Antarctic minke whale earplugs, and indicate that the bulla method is less precise than the earplug method. The results here were based on 6-12 parallel readings, and Olsen and Skaug (2002) indicated that 2-8 parallel readings were necessary to reach ±2 years precision. Conducting this many parallel readings on all whales sampled from the Norwegian catch would require a much larger effort than is currently employed.

Even though the precision was reasonable, the analysis of bias did not follow the expectations illustrated in Fig. 1. In Fig. 5 the expected increase in ovulations with age was not found. With a R^2 of 0.0014 and whales as young as seven years having 18 ovarian scars, Fig. 5 indicated that the age estimates were underestimates of the true age. The plot of ovulations versus age in Fig. 1 is simplistic, ignoring age-specific ovulation rates. However, with no indication of change in ovulation rate of North Atlantic minke whales, the model is not unrealistic, but more research should be conducted on evaluating age-specific ovulation rates for the species. Fig. 1 shows that even with a density-dependent shift in ASM and ovulation rates varying from 0.5-2 per year an increase in the numbers of ovulations with increasing age would be expected if the age estimates were unbiased. Fig. 5 showed no such relationship, which leaves the conclusion that bulla age estimates are not representative of the true age, and are severely biased. The steady downward trend in the residuals (Fig. 6) further strengthens this conclusion, and high positive residuals for the youngest animals show that these were the most problematic to age. None of the modelled ages exceeded 15 years, which is low compared with Antarctic minke whales who have been shown to live for more than 30 years (Kato *et al.*, 1991a). Studies by Christensen (1981) and Olsen and Sunde (2002) also indicated that North Atlantic minke whales could live for more than 30 years, and that ages exceeding 15 years were common. This shows that it becomes progressively more difficult to detect the outermost GLGs of older animals. However, excluding animals with a modelled age > 10 years would have little effect on the bias observed in Fig. 5, as Fig. 6 indicates a steady downward trend in bias through all age-classes. Some of the variability found in Fig. 5 may be due to variations in ovulation rate within the population, and individual differences in ASM. However, even with large variability in ovulation rate and ASM as explored in Fig. 1 some correlation between age and ovulations should be expected. Assuming that the model of one ovulation per year and ASM of eight years is correct, the residuals indicated an average underestimation of 37%, but ranging from an overestimation of 28% to an underestimation of 72%. However, without any known-age whales it is impossible to determine the exact magnitude of the bias.

Previous results on sexual dimorphism in length published by Christensen (1975), Jonsgård (1951), Larsen and Kapel (1982), Larsen and Kapel (1983), and Fig. 3 all show that females grow to larger sizes than males. When the von Bertalanffy growth model failed to achieve the same results (Fig. 7, Fig. 8) it was due to the clumped distribution of the data, the large variability in estimated age for a given length and the lack small animals which are crucial in fitting a von Bertalanffy curve correctly. The large difference in the k and t_0 parameters between the sexes (Table 2) also indicated that fitting the growth model was not successful. Even though there is sexual dimorphism in maximum body length, the length at birth and growth rate between the sexes should be very similar, and at least not as divergent as indicated from these analyses (Table 2).

The analysis of bias in this paper was based on an indirect approach since known-age animals or a validated ageing method for this species are missing. Therefore, the number of ovulations and body length were used as controls. These two parameters can themselves be subject to error, for instance the number of ovulations was determined from one examination of the ovaries by one reader. Few inconsistencies in their interpretation were expected, but a parallel ovary-examination study should be undertaken to verify this. Body length on the other hand has less potential error, but since the whales were measured at sea with head and tail sticking outside each side of the boat this might add some random error to the length measurements.

Reader variability was to a large extent the cause of the low precision (Olsen and Skaug, 2002), but the high bias is probably due to a combination of reader and biological effects. These biological effects are most probably the inconsistent nature of bulla GLG formation in the periosteal layer. The underestimation bias identified in the present analysis show that readable GLGs are not always formed in bulla every year. Adding to this basic uncertainty of GLG formation, the bulla GLGs observed were not continuous through a whole bulla section, but rather found at the peaks on small ridges or in cracks of the bulla. It is difficult to follow a single GLG through a single bulla segment, and almost impossible to find the same GLG in a segment cut only a few centimetres to either side of the first segment. In addition, many bulla GLGs are often very faint, thin and difficult to detect, while others are broad and clear. In some segments, a combination of thin and broad GLGs have been found making interpretation almost impossible. Bone resorption in bulla occurs along the mesosteal-periosteal interface (Christensen, 1981), and cannot explain the variable GLG widths and faintness observed in the outer parts of the periosteal layer. It does however explain why it is often difficult to identify the neonatal line in bulla segments. Adding to this confusion is the difficulty of determining where the neonatal line is and where post-parturition growth begins. These biological quirks makes a large degree of the bulla age determination procedure dependent on the subjective decisions of the reader and explains the lower precision and high bias of the bulla method compared to the earplug method.

From a management objective, the problem of precision in the bulla method is solvable through increasing the number of parallel readings of bulla segments from the same animal. However, the high bias will greatly limit any age-based monitoring of the North Atlantic minke whale population. The bias will tend to compress and level out even

determinable age-classes (Bradford, 1991), making it impossible to monitor changes in age-dependent demographic parameters. It seems impossible to circumvent this bias in any way, and therefore other age determination methods, such as the aspartic acid racemisation technique (Olsen and Sunde, 2002) should be employed to age North Atlantic minke whales.

ACKNOWLEDGEMENTS

I wish to thank Arild Leithe and Bjørn Bergfløtdt for reading all the bulla segments, Hans J. Skaug for the GLMM modelling, Leif Nøttestad and Nils Øien for reviewing previous versions of the manuscript, and the crew and owners of M/S *Feie* and M/S *Kato* for cooperation and help during the sample collection. The comments of two reviewers are appreciated. This project was funded as part of Norwegian Research Council project number 127211/120.

REFERENCES

Birney, E.C., Jenness, R. and Baird, D.D. 1975. Eye lens proteins as criteria of age in cotton rats. *J. Wildl. Manage.* 39(4):718-28.

Bowering, W.R. and Nedraas, K.H. 201. Age validation and growth of Greenland halibut (*Reinhardtius hippoglossoides*): A comparison of populations in the northwest and northeast Atlantic. *Sarsia* 86:53-68.

Bradford, M.J. 1991. Effects of ageing errors on recruitment time series estimated from sequential population analysis. *Can. J. Fish. Aquat. Sci.* 48:555-8.

Brody, S. 1945. *Bioenergetics and Growth*. Reinhold Publishing Co., New York.

Christensen, I. 1975. Preliminary report on the Norwegian fishery for small whales: expansion of Norwegian whaling to Arctic and northwest Atlantic waters, and Norwegian investigations of the biology of small whales. *J. Fish. Res. Bd Can.* 32(7):1083-94.

Christensen, I. 1981. Age determination of minke whales, *Balaenoptera acutorostrata*, from laminated structures in the tympanic bullae. *Rep. int. Whal. Commn* 31:245-53.

Christensen, I. 1992. Age determination of baleen whales. *North Atl. Stud.* 2:32-5.

Christensen, I. 1995. Interpretation of growth layers in the periosteal zone of tympanic bullae from minke whales, *Balaenoptera acutorostrata*. pp. 413-23. *In:* A.S. Blix, L. Walløe and O. Ulltang (eds.) *Whales, Seals, Fish and Man. Proceedings of the International Symposium on the Biology of Marine Mammals in the Northeast Atlantic*. Elsevier, Amsterdam. 720pp.

Christensen, I., Haug, T. and Wiig, O. 1990. Morphometric comparison of minke whales *Balaenoptera acutorostrata* from different areas of the North Atlantic. *Mar. Mammal Sci.* 6(4):327-38.

George, J.C., Bada, J., Zeh, J., Scott, L., Brown, S.E., O'Hara, T. and Suydam, R. 1999. Age and growth estimates of bowhead whales (*Balaena mysticetus*) via aspartic racemization. *Can. J. Zool.* 77:571-80.

International Whaling Commission. 1994. Report of the Scientific Committee, Annex K. Area definitions for RMP implementations. *Rep. int. Whal. Commn* 44:175-6.

International Whaling Commission. 1999. International Whaling Commission Report 1997-98. *Ann. Rep. Int. Whaling Comm.* 1998:1-2.

Jonsgård, Å. 1951. Studies on the little piked whale or minke whale (*Balaenoptera acutorostrata* Lacépède). Report on Norwegian investigations carried out in the years 1943-1950. *Norsk Hvalfangsttid.* 40:209-32.

Kato, H., Fujise, Y. and Kishino, H. 1991a. Age structure and segregation of southern minke whales by the data obtained during Japanese research take in 1988/89. *Rep. int. Whal. Commn* 41:287-92.

Kato, H., Zenitani, R. and Nakamura, T. 1991b. Inter-reader calibration in age reading of earplug from southern minke whales, with some notes of age readability. *Rep. int. Whal. Commn* 41:339-43.

Larsen, F. and Kapel, F.O. 1982. Norwegian minke whaling off West Greenland, 1976-80 and biological studies of West Greenland minke whales. *Rep. int. Whal. Commn* 32:263-74.

Larsen, F. and Kapel, F.O. 1983. Further biological studies of the West Greenland minke whale. *Rep. int. Whal. Commn* 33:329-32.

Laws, R.M. 1958. Recent investigations on fin whale ovaries. *Norsk Hvalfangsttid.* 47(5):225-354. [In Norwegian and English].

Morris, P. 1972. A review of mammalian age determination methods. *Mammal Rev.* 2(3):69-104.

Nuckle, J. and Bergeron, J.-M. 1983. Eye lens weight as a criterion for age classification of wild woodchucks. *J. Wildl. Manage.* 47(3):846-50.

Ohsumi, S. and Masaki, Y. 1975. Biological parameters of the Antarctic minke whale at the virginal population level. *J. Fish. Res. Bd Can.* 32(7):995-1004.

Olsen, E. 1997. A study of the reproductive potential of North Atlantic minke whales (*Balaenoptera acutorostrata*, Lacépède 1804) in the period 1972-79. Master Thesis, University of Bergen, Bergen, Norway. 76pp.

Olsen, E. and Skaug, H.J. 2002. Methods for variance decomposition of age estimates applied to age estimates of North Atlantic minke whales. 19pp. In: E. Olsen. *Age Determinations of Minke Whales*. Ph.D. Thesis, University of Bergen, Bergen, Norway. 184pp.

Olsen, E. and Sunde, J. 2002. Age determination of minke whales (*Balaenoptera acutorostrata*) using aspartic acid recemization technique. *Sarsia*: 87:1-8.

Richards, L.J., Schnute, J.T., Kronlund, A.R. and Beamish, R.J. 1992. Statistical models for the analysis of ageing error. *Can. J. Fish. Aquat. Sci.* 49:1801-15.

Schnute, J.T. and Richards, L.J. 1995. The influence of error on population estimates from catch-age models. *Can. J. Fish. Aquat. Sci.* 52:2063-77.

Sukhovskaya, L.I., Klevezal, G.A., Borisov, V.I. and Lagerev, S.J. 1985. Use of bone layers to determine age in minke whales. *Acta Ther. Sinica* 30(17):275-86.

Factors affecting the precision of age determination of sperm whales (*Physeter macrocephalus*)

KAREN EVANS*, MARK A. HINDELL*, KELLY ROBERTSON#, CHRISTINA LOCKYER+ AND DALE RICE**

Contact e-mail: Karen.Evans@csiro.au

ABSTRACT

Teeth from 92 sperm whales were prepared by etching for age determination. The total number of growth layer groups (GLGs) in the dentine of each tooth was determined from three to five reading sessions by a single reader. Four other readers, as part of a cross-reading experiment, read a subset of these teeth ($n = 5$). This study investigated: (1) intra- and (2) inter-reader precision in GLG counts; (3) possible variation in growth structure deposition between different teeth within the same individual; (4) the use of photographs to identify and count GLGs and the effect of this technique on the precision of counts; and (5) mineralisation anomalies in tooth sections and the possible effects these may have on GLG count precision. Intra- and inter-reader precision was determined using coefficients of variation (CV) and indices of precision (D). Total numbers of GLGs estimated from individual teeth ranged from 0.75-64 ($\bar{x} = 32.8$, $n = 92$). Intra-reader mean CV was 10.6 and mean D was 4.8. Inter-reader mean CV ranged from 4.8-12.3 and mean D ranged from 2.8-7.1. Differences in final counts between readers appeared to be the result of differing interpretation of GLGs and this was the largest factor affecting the precision of GLG counts. While GLG counts between teeth in the same individual varied, it is possible that this variation was due to within reader variation rather than variation in the development of growth structures, but establishment of this cause is confounded by differential tooth wear. Use of photographs increased the definition of growth structures, decreasing the variation between GLG counts within reading sessions. The incidence of mineralisation anomalies and the closure of the pulp cavity increased with increasing GLG counts in individuals, but were not consistent between teeth from the same individual. These factors, while potentially affecting the accuracy of GLG counts in relation to age estimates, had little effect on the precision of GLG counts. The lack of an ability to validate age estimates in this species and the large inter-reader variation seen in this study suggests that age estimates based on GLG counts in this species are subjective and can only be regarded as relative. High-quality photographs of tooth sections should be used to verify GLG counts with other readers, resulting in 'consensus counts' generated by a number of readers, ensuring interpretation of the same structures and confidence in comparing GLG counts produced in different studies.

KEYWORDS: AGE DETERMINATION; SPERM WHALE; AUSTRALASIA; SOUTHERN HEMISPHERE; STRANDINGS

INTRODUCTION

The determination of the age of animals is important in establishing the life history traits of individuals and populations. Integral to this is the development of an accurate age determination technique and the minimisation of any associated biases.

Growth layer groups (GLGs) in the teeth of sperm whales (*Physeter macrocephalus*) have been used to determine the age of individuals since the 1950s (Nishiwaki *et al.*, 1958; Ohsumi *et al.*, 1963; Gambell, 1977; Rice *et al.*, 1986). However, validation of the assumption that these GLGs are annual depositions, as is the case in most other marine mammals, has proven difficult. Validation techniques such as the use of 'known-age' individuals (Hohn *et al.*, 1989; Hohn, 1990) and tetracycline marking experiments (Myrick *et al.*, 1984; 1988; Brodie *et al.*, 1990) used in other species have not been feasible in sperm whales because of their size and the inability to keep captive individuals. Only limited mark-recapture studies investigating the accumulation rate of growth layers and studies calibrating seasonal changes in the thickness of the most recently formed dentine layer have been conducted on this species. These studies suggest that GLGs are deposited annually (Ohsumi *et al.*, 1963; IWC, 1967; 1971; Best, 1969) and as a result, studies involving the age determination of this species assume that each GLG represents one year's growth (Ohsumi, 1971; 1977; Lockyer, 1980; Rice *et al.*, 1986).

Another important concern associated with the aging of individual animals is that of the precision of counts of GLGs and, therefore age estimates (that is, the closeness of repeated GLG counts for the same individual). If final age estimates are the result of averaging the GLG counts from a number of reading sessions, the precision of GLG counts may have a major effect on the accuracy (the nearness of the final age estimate or GLG count to the actual age or number of GLGs) of the final estimate. As age increases, the pulp cavity in the tooth of a sperm whale fills in as a result of the deposition of further layers of dentine and eventually closes. Once the cavity is closed, the most recently deposited layers become compacted and are subsequently hard to discern. Mineralisation anomalies and dentinal resorption (Myrick, 1988; Lockyer, 1993) may also confuse the distinctiveness of GLGs, particularly in the recently deposited dentine of older animals, which may already be compromised by the closure of the pulp cavity. A number of publications have addressed variation in the accuracy of age determination from cetacean teeth associated with the preparation and reading techniques used (Anas, 1970; Hui, 1980; Hohn *et al.*, 1989; Hohn, 1990; Hohn and Fernandez, 1999). However, very few have addressed the problem of variation in precision (Donovan *et al.*, 1982; Mikhalev, 1982; Reilly *et al.*, 1983).

Variation in the number of GLGs in different teeth from the same individual may also be another source of bias in age determinations. Nishiwaki *et al.* (1958) found that teeth from

* *Antarctic Wildlife Research Unit, School of Zoology, University of Tasmania, GPO Box 252-05, Hobart, Tasmania, Australia, 7001. (*Present address: *Tropical and Pelagic Ecosystems, CSIRO Marine Research, GPO Box 1538, Hobart, Tasmania 7001, Australia).*
# *Southwest Fisheries Science Center, National Marine Fisheries Service, NOAA, PO Box 271, La Jolla, California, 92038, USA.*
+ *Danish Institute for Fisheries Research, Department of Fisheries, Charlottenlund Slot, Charlottenlund, DK2920, Denmark.*
** *National Marine Mammal Laboratory, National Marine Fisheries Service, NOAA, 7600 Sand Point Way, Seattle, Washington, 98115, USA.*

both the mandibular and maxillary jaw from the same individual in sperm whales contained similar numbers of growth layers. Conversely, the Workshop on Age Determination of Odontocete Cetaceans and Sirenians found that the number of GLGs varied between different teeth from one individual (Perrin and Myrick, 1980). However, the dataset was not large enough to test this statistically and it was recommended that the number of GLGs in complete series of teeth from both the mandibular and maxillary jaw of a number of individuals of varying ages be assessed. Bottlenose dolphins (*Tursiops truncatus*) were found to contain different numbers of GLGs in different teeth from the same individual (Hui, 1980), possibly because teeth in the anterior of the jaw ceased depositing dentine after 10 to 12 GLGs and posterior teeth ceased deposition of dentine at any time after 15 GLGs. However, both Myrick (1988) and Lockyer (1993) found that different teeth from the same individuals in spinner (*Stenella longirostris*), pantropical spotted (*Stenella attenuata*), common (*Delphinus delphis*) and bottlenose dolphins and long-finned pilot whales (*Globicephala melas*) showed similar growth patterns yielding similar age estimates (counts in *G. melas* differed by 0 to 4 GLGs).

Three mass strandings of sperm whales on the west and northwest coasts of Tasmania, Australia in 1998 provided material with which these problems could be investigated. This paper presents the results of investigations into variations in age estimates (1) within and (2) between readers; (3) between different reading methods; (4) between different teeth derived from the same individual; and (5) in relation to tooth morphology.

MATERIALS AND METHODS

Preparation of teeth

Near-complete or mid-sections of lower jaws with teeth were collected from 92 sperm whales involved in three mass strandings on the north and west coasts of Tasmania in 1998 (STR1: Ocean Beach, Strahan, $n = 56$; STR2: Greens Pt Beach, Marrawah, $n = 29$; STR3: Black River Beach, Stanley, $n = 7$). Other aspects of these strandings are reported in Evans *et al.* (2002). The least worn and straightest first or anterior-most mandibular tooth from each individual was sectioned along the bucco-lingual plane, and one half-section polished and then etched in 15% formic acid until clear, easily discernible dentinal layers or growth layer groups (GLGs) were produced. Teeth from calves were thin-sectioned, stained and mounted on microscope slides. Details of these methods are given in Evans and Robertson (2001).

Age determination

The total number of GLGs in each of the 92 tooth sections was determined three times per session in three ($n = 3$), four ($n = 7$) or five ($n = 82$) sessions by a single reader (KE). The number of reading sessions was determined by the variability of GLG counts. For those teeth for which counts were not repeatable or at least two of the three counts were not close (within ± 2GLGs), counts were repeated an additional one or two times. Time intervals between the sessions varied from seven to 92 days. Each reading was made without reference to previous readings or additional information on individuals (e.g. size, sex) and teeth were read in random order during each session.

Growth layer groups were interpreted as those identified in the report of the Workshop on Age Determination of Odontocete Cetaceans and Sirenians (Perrin and Myrick, 1980) as 'a repeating or semi-repeating pattern of adjacent groups of incremental growth layers within the dentine which is defined as a countable unit involving a change from a ridge to groove' in the case of etched teeth and 'intensely stained to lightly stained' in the case of thin-sectioned, stained teeth. For those specimens in which the neonatal line could be identified (many of the teeth had worn tips and therefore were missing the enamel, neonatal line and the first few GLGs), this was not included in the total number of GLGs.

The final age estimate for each individual was determined as either the most repeated GLG count (all session estimates pooled) or where there was no repeatability of counts ($n = 28$) the mean of all counts. It was assumed that skill in reading GLGs and estimating age increased with reader experience. To determine whether this had an effect on counts, a two-way ANOVA (with session number and tooth section as independent variables) was conducted on GLG counts to determine whether counts differed significantly between sessions.

Assessment of intra-reader variation

For each tooth the standard deviation was calculated from all counts (the three counts from each of the three, four or five sessions all pooled). Following Chang (1982) and Reilly *et al.* (1983), the coefficient of variation (the standard deviation as a fraction of the mean expressed as a percentage: $SDx100/\bar{x}$) and an index of precision (the percent error contributed by each observation to the average age class: $D = CV/\sqrt{n}$) were calculated. The CV and D were plotted against GLG counts to determine if there was any effect of the number of GLGs (and therefore age) on the precision of counts.

Assessment of inter-reader variation

A sub-sample containing five of the original 92 tooth sections and an associated photograph of each were supplied to four additional readers for comparative GLG determination. All readers had previous experience in counting GLGs from sperm whale teeth. No information on the animal or the stranding from which it was derived was supplied to the readers. Readers were supplied with a standard data form and were requested to estimate the number of GLGs directly from each tooth at least three times with a minimum of five to seven days between reading sessions and without reference to previous readings. Each reader was also requested to mark on the associated photograph of each tooth what they had interpreted and counted as GLGs, in an effort to establish areas in which variation, if it existed, occurred.

Individual reader CV and D were calculated for each tooth to quantify individual reader precision. Actual counts from individual teeth were compared between readers using a two-way ANOVA (with reader and tooth section as independent variables). Where this test revealed that there were significant differences in GLG counts between readers, the relevant photographs on which each reader had mapped their interpretation of GLGs were studied and any differences in the definition of GLGs noted.

Assessment of counts from different teeth from the same animal

For seven whales, an additional 13 teeth were prepared for age estimation (providing a total of seven teeth from each side of the jaw and a total of 14 for each animal). These animals were selected randomly from a subset of the original that contained animals from which more than seven teeth on

each side of the jaw had been collected. The teeth selected were dependent on the number of teeth collected from the jaw. Where more than seven teeth from either jaw were collected, teeth were selected evenly along the length of the jaw. In all cases teeth from matching positions on both sides of the jaw were used. The number of GLGs in each tooth was estimated using the methods detailed above without reference to other teeth from each individual. GLG counts derived from teeth on the left and right sides of the jaw in an individual were compared for differences using a paired t-test.

To determine whether the numbers of GLGs in the 14 teeth of an individual were significantly different, an ANOVA with a Tukey HSD pairwise comparison was used. For each tooth the standard deviation, CV and D were calculated from all counts. To determine whether the number of GLGs did in fact vary between teeth in each animal, it was necessary to separate that variation associated with the reader from true differences in the number of GLGs present in each tooth. D values for each tooth from an animal were plotted with the mean D calculated from the assessment of within-reader variation (the mean overall D). Where D values for each tooth were lower than the mean overall D, any variation in GLG counts were regarded as true variation in the number of GLGs. Where D values for each tooth were the same or higher than the mean overall D, variation in GLG counts were regarded as a factor of reader variation. A one-way t-test was used to test for the presence of such differences.

Assessment of direct tooth counts vs photo counts

All teeth prepared (for both estimates of age of individuals in each stranding and for comparative counts of different teeth from the same individual) were digitally photographed ($n = 171$). The number of GLGs in each tooth section was determined from these images. A paired t-test was used to compare GLG counts derived from photos against final GLG counts derived from direct readings. To determine if there was any difference in the precision of counts between this method and that from counts taken directly from teeth, a sub-sample of 50 randomly selected tooth images were read a further two times (for a total of three readings). Both CV and D were calculated for each method and then compared.

Tooth morphology

The number of pulp stones, the presence of mineralisation interferences (occlusions), and the state of the pulp cavity (whether it was open or closed) were determined for each tooth section. Anomalies were classified according to Lockyer (1993). The CV and D calculated during age determination were log-regressed against the state of the tooth cavity and against the presence of pulp stones and regressed against the number of pulp stones to determine whether tooth morphology factors effected CV and D.

RESULTS

Assessment of intra-reader variation

GLG counts from sperm whales in this study ranged from 0.75 to 64 GLGs (mean = 32.8±13.2, $n = 92$). There were no significant differences among GLG counts estimated in the five sessions (ANOVA, $F_{4,91} = 0.9$, $P = 0.5$). For those estimates where there was no consensus of GLG counts between sessions, 89.3% contained estimates that differed by one GLG and 96.4% contained estimates that differed by two GLGs.

The mean CV was 10.6±6.3 and mean D was 4.8. There was no significant relationship between CV or D and the number of GLGs (Regression, CV: $r^2 = 0.001$, $F_{1,90} = 0.1$, $P = 0.7$; D: $r^2 = 0.004$, $F_{1,90} = 0.4$, $P = 0.6$; Fig. 1).

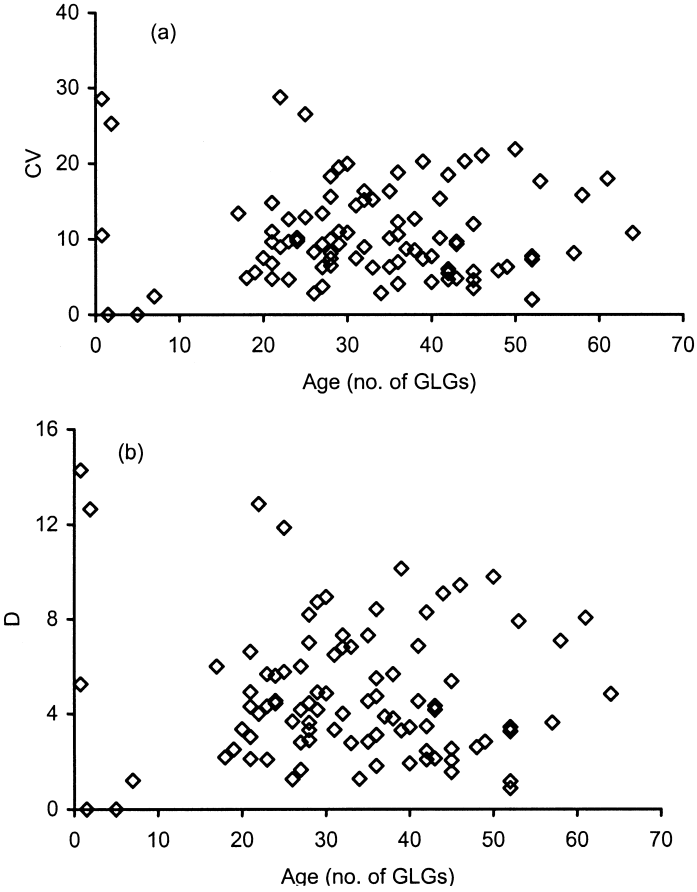


Fig. 1. CV and D calculated from estimated number of GLGs in teeth from sperm whales ($n = 92$): (a) CV; (b) D.

Assessment of inter-reader variation

The difference in the number of GLGs estimated for each of the five teeth ranged considerably between readers, from one to 21 GLGs (means: 5.0-11.8 GLGs; Tables 1 and 2), increasing with teeth from older animals (Fig. 2). GLG counts were found to be significantly different between readers (ANOVA, $F_{4,16} = 2.2$ $P = 0.02$). Mean CV ranged from 4.8-12.3 and mean D ranged from 2.8-7.1 across readers.

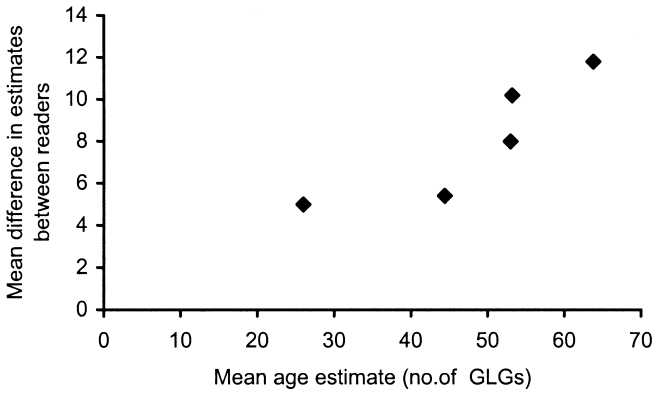


Fig. 2. Average difference in the estimates of the number of GLGs between five readers and the average age estimated for five sperm whale teeth.

Table 1

Estimated numbers of GLGs by five readers from three reading sessions of five sperm whale teeth.

Tooth	Number of estimated GLGs														
	Reader 1			Reader 2			Reader 3			Reader 4			Reader 5		
001	41	39	44	24	39	39	52	54	46	54	44	48	41	43	42
002	53	51	51	43	47	49	64	57	71	64	62	58	43	43	43
003	24	21	24	23	27	20	35	33	32	29	28	29	21	23	21
004	39	54	47	50	44	56	65	56	60	61	65	62	44	51	47
005	65	68	68	63	60	61	69	65	80	75	70	70	45	53	52

Table 2

Final estimates of GLG numbers for five sperm whale teeth by five readers and mean standard deviation, mean CV and mean D.

Tooth	Final estimate of the number of GLGs				
	Reader 1	Reader 2	Reader 3	Reader 4	Reader 5
001	41	39	51	49	42
002	52	46	64	61	43
003	24	23	33	29	21
004	45	50	60	63	47
005	65	61	71	70	52
Mean SD	3.2	4.6	5.0	2.7	2.0
Mean CV	7.6	12.3	8.4	4.9	4.8
Mean D	2.8	7.1	4.9	2.9	2.8

Assessment of counts from different teeth derived from the same individual

No significant differences were found in the number of GLGs, the CV or D between teeth in the left and right jaws of any of the individuals in the dataset. However, significant differences were found between GLG counts from teeth in different positions along the tooth row within an individual in six of the seven animals (Tables 3 and 4). When D values calculated from readings of each tooth in each individual were compared with the mean overall D, significant differences were found in only one individual (t-test, $t_{13} = 4.9$, $P < 0.001$; Fig. 3). GLG counts from different teeth in this animal ranged from 18-33. GLG counts varied increasingly with age (Fig. 4).

Table 3

Minimum, maximum and mean number of GLGs, mean SD, CV and D estimated from multiple teeth (n=14) of seven sperm whales at STR2.

	Whale number						
	1	4	8	16	21	25	27
Min. no. GLGs	19	17	28	18	20	25	39
Max. no. GLGs	52	25	43	37	32	43	64
Mean no. GLGs	30.9	20.6	34.2	27.8	25.3	33.4	44.8
Mean SD	3.5	2.9	4.0	3.6	3.7	4.0	3.6
Mean CV	11.9	13.4	11.2	13.3	14.6	13.0	8.4
Mean D	5.4	6.0	5.1	6.0	6.5	6.1	3.7

Table 4

Results of ANOVAs conducted on GLG counts from teeth (n=14) in the jaws of seven sperm whales at STR2.

	Whale number						
	1	4	8	16	21	25	27
df	13	13	13	13	13	13	13
F-ratio	19.6	1.7	3.0	9.6	2.1	10.8	7.6
P	<0.001	0.09	0.002	<0.001	0.03	<0.001	<0.001

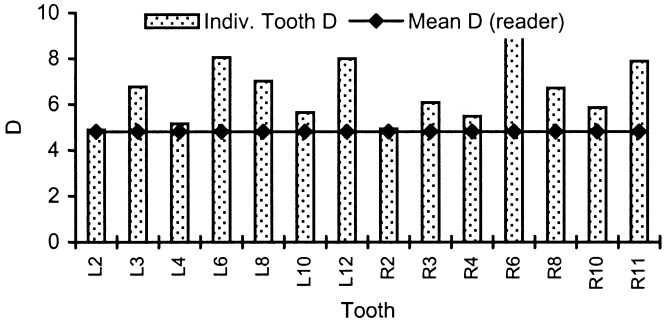


Fig. 3. D values for teeth (n = 14) in STR2(21) and the mean overall reader D.

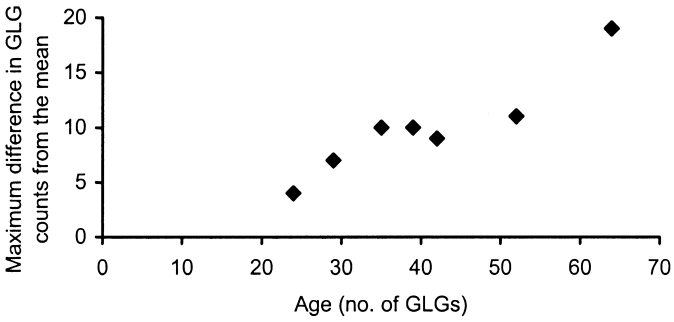


Fig. 4. Maximum difference in age estimates from the mean of age estimates derived from sets of 14 teeth prepared and examined in seven sperm whales.

Assessment of direct tooth counts vs. photo counts

GLG counts derived from photographs were only available for four of the five readers. Counts derived from photographs of individual teeth were significantly higher ($\bar{x}$ = 36.1±10.7 GLGs) than those derived from direct examination ($\bar{x}$ = 32.9±9.5) of teeth (t-test, $t_{170} = 9.8$, $P < 0.001$). The mean difference in GLG counts between these two methods was 3.2 GLGs.

The mean CV of the subset of photographs that were read several times was 8.0±3.9 and the mean D was 4.6±2.3, while the mean CV derived from direct counts of these teeth was 11.8±5.7 and the mean D 5.4±2.6.

Tooth morphology

The mean number of GLGs in teeth with pulp stones was 36.2±11.5 (range: 5-64 GLGs, n = 67) and mean number of pulp stones present was 6.4±9.7 (range: 0-53, n = 92). Both the presence and number of pulp stones in tooth sections were significantly related to the number of GLGs (Presence: log-regression, $t_1 = 3.6$, $P < 0.001$; Number: regression, $r^2 = 0.04$, $F_{1,90} = 4.4$, $P = 0.04$; Fig. 5a). In those individuals where multiple teeth were examined, neither the presence nor the numbers of pulp stones were constant throughout different teeth (Fig. 6). The maximum range in pulp stone number between teeth in an individual was 0-32. There was no significant relationship between either CV or D and pulp stone presence or number (Figs 5b and 5c).

The incidence of a closed pulp cavity is related to increasing age. GLG counts from animals in which the tooth examined had an open pulp cavity were significantly lower (26.4±10.4) than those of animals in which the tooth examined had a closed (45.5±8.0) pulp cavity (t-test, $t_{30} = -8.4$, $P < 0.001$). In six of the seven individuals where multiple teeth were examined, the state of the pulp cavity was not consistent along the tooth row; instead each contained a mixture of teeth with open cavities and closed

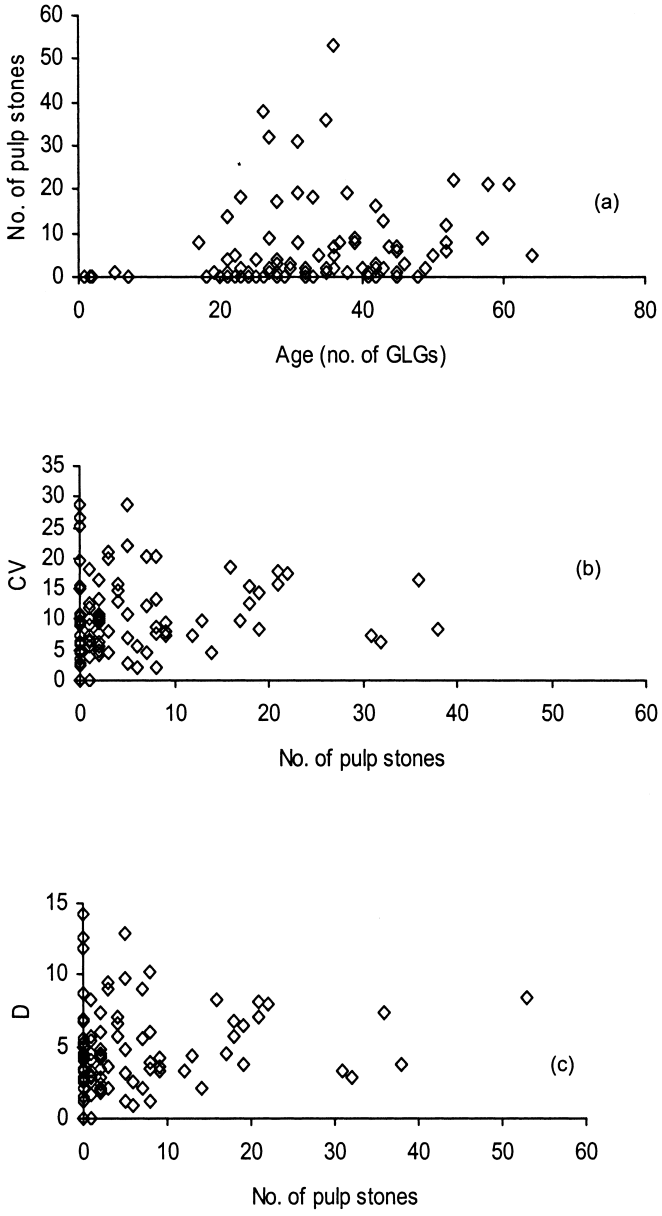


Fig. 5. Number of pulp stones present and the estimated age, CV and D calculated from sperm whale tooth sections ($n = 92$): (a) age; (b) CV; (c) D.

Fig. 6. Differences in mineralisation anomalies and state of pulp cavity between two teeth from STR2(25): (a) L1; (b) L11.

cavities (e.g. Fig. 6). The number of teeth in the tooth row with closed pulp cavities increased with the number of GLGs. CV and D were not significantly related to the closure of the pulp cavity in tooth sections (log-regression, CV: $t_1 = 0.004$, $P = 0.1$; D: $t_1 = 0.01$, $P = 0.1$).

The mean number of GLGs in teeth containing occlusions was 41.0±8.6 (range: 30-61, $n = 11$). Occlusions were not common throughout teeth from the same individual. Of two individuals where multiple teeth were examined and occlusions were present, one of 14 teeth contained an occlusion in one and four out of 14 teeth contained an occlusion in the other.

DISCUSSION

Assessment of intra- and inter-reader variation

Average CV and D calculated for intra-reader variation are similar to those presented in Reilly *et al.* (1983) for pantropical spotted dolphins and suggest that GLG counts from this dataset were relatively precise. Unlike in other studies (Doubleday and Bowen, 1980; Reilly *et al.*, 1983; Bjørge *et al.*, 1995), this degree of precision did not decrease with increasing animal age. GLG counts did not appear to vary across reading sessions either, with no significant differences between session estimates. This suggests, at least in this study, that the precision of GLG counts was relatively constant throughout the age determination exercise.

However, GLG counts and average CV and D varied substantially between readers (by up to 21 GLGs) and this variation increased with increasing GLG number, although again, values for CV (4.8-12.3) and D (2.8-7.1) were similar to or lower than those calculated in other studies. Mean D values in Reilly *et al.* (1983) ranged between 2.8 and 6.6, while those in Chang (1982) ranged from 3.4-9.8. This variation has been found to increase with increasing specimen age in a number of other cetacean species (Reilly *et al.*, 1983; Bjørge *et al.*, 1995; Hohn and Fernandez, 1999) and is due to a decreasing ability to interpret growth structures in older animals. The deposition of growth layers becomes more highly compacted as the pulp cavity area fills in and its size decreases, making it harder to discern individual GLGs from one another.

None of the readings from direct examination of teeth coincided between readers, although in all at least two of the readings varied by less than three GLG. IWC (1969) reported that the average deviations from the mean of the age estimate ranged from +4.5 to –3.1 for 11 readers examining

the same teeth, although estimates for eight of the 11 readers were ±1 GLG. Donovan *et al.* (1982) reported a significant difference (using Friedman's test) between the 'best' estimates of six readers when reading 50 etched teeth but no significant difference was found when four of the six readers (from the same 'school' of reading) were compared. When the age estimates of the remaining two readers were compared to the others, the average deviations from the mean were +1.42 and –1.76[1]. Mikhalev (1982) added the age estimates of two further readers experienced in reading sperm whale teeth to the results of Donovan *et al*. (1982) and observed further variation. However, he also noted that the average difference between the 'extreme' readers was only 3.2 GLGs; the maximum was 10 GLGs (in two teeth). These results all reveal a degree of subjectivity in the interpretation of growth layers in teeth. The implications of this subjectivity, particularly relevant for inter-study comparisons, depend on the use to which the age data are put and the way in which the 'best' estimate is arrived at. In this regard, the question of how to deal with worn teeth is important.

Examination of the associated photographs for the cross-reading experiment also highlighted this subjectivity, demonstrating that differences in GLG counts were due to differences in the interpretation of GLGs (i.e. what were regarded as accessory layers by one reader were regarded as GLGs by another; Fig. 7). This has substantial implications when comparing age estimates between studies. Attempts to standardise the definition and interpretation of GLGs in age determination studies were made during the International Whaling Commission's Workshop on Age Determination in Cetaceans and Sirenians. While the report of this workshop was published (Perrin and Myrick, 1980) and a number of papers (Nishiwaki *et al.*, 1958; Ohsumi *et al.*, 1963; Best, 1969; Scheffer and Myrick, 1980) have provided photographs of sectioned teeth illustrating GLGs (as defined by the authors), no quantitative and objective method to assist researchers in the laboratory has yet been published. Definitions of GLGs depend, as a result, on the interpretation

(a) Reader 1. Estimated number of GLGs: 57.

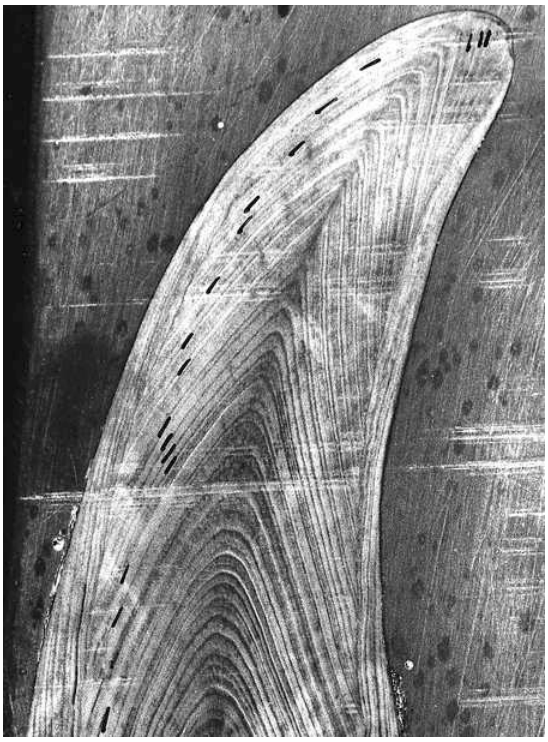


Fig. 7a. Growth layer groups in a tooth from SPW2(25) as interpreted by three readers.

[1] For comparisons with present study, the mean CV for all readers was 7.55 and mean D was 1.26; for the four readers from the same school the values were 4.97 and 0.83, respectively (Donovan, pers. comm.).

(b) Reader 2. Estimated number of GLGs: 50.

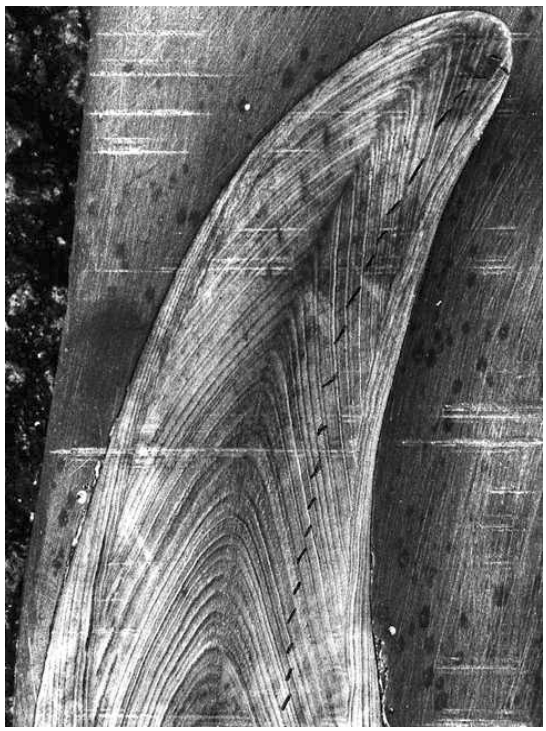


Fig. 7b.

(c) Reader 3. Estimated number of GLGs: 61.

Fig. 7c.

of the individual or the laboratory at which age estimates are being determined and are therefore qualitative and subjective.

Assessment of counts from different teeth from the same individual

GLG counts from different teeth from individual whales differed significantly in six out of seven specimens. However, of these, the mean D generated from each set of 14 teeth was not significantly different from the mean D of the reader in all but one case, suggesting that variation in GLG counts between teeth was less likely to be the result of differences in the growth structures between teeth and may have been associated with intra-reader variation. The variation in GLG counts from teeth in the individual for which tooth D values differed significantly from the mean reader D could also be explained similarly. The mean D value generated from this individual was 6.5, a value higher than that of the overall mean reader D (4.81). However, attributing these differences to variation in the precision of GLG counts between teeth is confounded by the effects of differential tooth wear. Hui (1980) found that anterior teeth in bottlenose dolphins yielded lower numbers of GLGs than posterior teeth. Rather than these teeth containing varying growth structures, differential use and therefore differential wearing of these teeth may have resulted in varying GLG counts. If differential wear does have a significant impact on the determination of GLG counts in different teeth from the same individual, determining whether growth structure variation does occur between teeth becomes difficult.

Reducing intra-reader variation and thereby increasing the precision of GLG counts as well as devising some means by which tooth wear could be quantified would assist in establishing the source of this variation in GLG counts between teeth from the same individual. A larger formal internationally organised trial should be considered to quantify such factors and establish means by which these can be calibrated across studies.

Assessment of direct tooth counts vs photo counts

Differences of up to 21 GLGs between readers have serious implications for the validity of comparative studies, particularly in this long-lived species where there are no real means of verifying ages (e.g. via known-age animals or tetracycline experiments). However, the use of photographic techniques in the determination of GLG counts may serve to reduce this variation. Both the overall variation in estimates relative to the mean (CV) and the error contributed by each observation (D) decreased in GLG counts derived from the multiple readings of photographs in comparison to those derived from direct counts. The higher counts produced by all readers using photographs may be the result of two factors: (1) less confusion in interpreting between GLGs and accessory layers or (2) greater clarity of and contrast in growth structures causing accessory layers to appear as substantial a growth structure as GLGs. When counting GLGs, the reader must make a decision as to whether a growth structure is a GLG or an accessory layer. Readers may either be cautious, only interpreting the most clear structures as GLGs (and thereby perhaps underestimating the true number of GLGs) or may interpret most growth structures as GLGs (possibly including the clearest and most highly contrasted accessory layers as GLGs). This interpretation is highly subjective, but the fact that all readers counts increased while the individual reader CV values decreased when using photographs suggests that the photographs resulted in the same effect on reader interpretation of growth structures and overall increased reader precision.

Hohn (1980) found that in comparing the use of polarised light, microradiography and scanning electron microscopy in age estimation techniques, scanning electron microscopy provided images in which GLGs were easiest to read. This was attributed to the higher contrast in topographic relief between the layers of each GLG. Bow and Purdy (1966) also found that the use of photographs of etched teeth increased contrast and maximised shadow detail between growth layer groups with the end result of decreasing errors in counts. While only the effect of the use of high quality photographs on GLG counts was studied here, other photographic methods such as the use of 3-D stereographic techniques should be considered in efforts to increase the clarity of and the contrast between individual GLGs and between GLGs and accessory layers, thereby increasing reader precision.

Tooth morphology

Mineralisation anomalies such as pulp stones and occlusion events have been documented in cetaceans on numerous occasions (Klevezal and Myrick, 1984; Myrick, 1988; Lockyer, 1993; 1995), but no assessment has been made on the effect of such anomalies on age estimation. Pulp stones are discrete events within the dentine of tooth sections, in most instances having little effect on the appearance of GLGs. Large pulp stones can bend GLGs, or may obscure that part of the GLG situated in the area of the pulp stone. Regardless of pulp stone size, GLGs can still be identified in the dentine of tooth sections. As a result, it would be expected that such events would have little effect on the precision (as found here) or the accuracy of GLG counts. Occlusions however, may obscure GLGs by disrupting lamina formation to the extent that they are no longer clearly defined. This may not affect the precision of GLG counts, since the same number of laminae actually defined within and outside the mineralisation interference area can be identified. However, such events have implications for the accuracy of GLG counts, especially in older animals in which both the incidence and the number of mineralisation anomalies are higher. Similarly, closure of the pulp cavity and the subsequent compacting and obscuring of GLGs is less likely to affect the precision of GLG counts (as found here), but is likely to affect the accuracy of GLG counts.

Even for the same individual, the presence and extent of mineralisation anomalies and the closure of the pulp cavity in differing teeth can be highly variable. Pulp stones form in the pulp and may not necessarily be incorporated into the dentine, or may spend varying amounts of time in the pulp before deposition in the dentine (Lockyer, 1993). As a result, varying numbers and positions of pulp stones in teeth from the same individual, as in this study are likely to occur (Fig. 6). If possible, rather than collecting a particular tooth from the jaw of an animal, several teeth should be collected and age estimates determined from the tooth with the least wear, the most highly defined growth layer groups, the minimum extent of mineralisation anomalies and if possible with an open pulp cavity (ensuring that GLGs have not become obscured with the closure of the cavity).

GLG counts and as a result, age estimates in this species are determined by an individual reader's interpretation of growth structures in tooth sections. Therefore, the largest factor affecting the precision of age estimates of individual animals is inter-reader variation in this interpretation. GLG

counts generated by a single reader can only be regarded as relative and comparable within a study, because any error introduced by the reader can be assumed to be relatively consistent across all estimates. However, large inter-reader variation compromises the ability to compare GLG counts and therefore age estimates between studies, especially when no indication of the precision of those age estimates is given. While there are currently no accurate means of determining the number of annual growth layers in this species, attempts should be made to increase the precision of age estimates, both within and between studies, and to devise more objective means by which GLG counts and therefore age estimates can be generated. The use of high quality photographs or other photographic techniques enabling clearer definition of GLGs should be investigated further as they may assist by increasing both intra- and inter-reader precision. Such photographic techniques could be used to verify GLG counts with other readers, ensuring interpretation of the same structures and facilitating 'consensus counts' generated by a number of readers, thereby increasing confidence in comparing age estimates between studies. Further studies investigating possible variability in growth structures between teeth from individuals and those enabling the separation of the effects of reader variability and the effects of differential wear in teeth on this variability should be initiated. Greater collaboration between investigators working on studies requiring age estimation of this species should be encouraged and is essential if standardisation of growth structure interpretation is to be achieved.

ACKNOWLEDGEMENTS

We thank Margie Morrice, Deborah Thiele, John van den Hoff, Graham McKenzie and all the volunteers that helped with the collection of samples. The Parks and Wildlife Service, Tasmania, particularly Hans Wapstra and the School of Zoology, University of Tasmania are acknowledged for their help with logistics. The Australian Antarctic Division, particularly Harry Burton and Andrew Fleming and the Southwest Fisheries Science Center, California, USA, particularly Susan Chivers, kindly provided laboratory space and equipment for the preparation of samples. John Bannister and William Perrin provided valuable comments on earlier versions of the manuscript. This study was conducted under the Parks and Wildlife Service, Tasmania permit 97/16 and was funded by Environment Australia.

REFERENCES

Anas, R.E. 1970. Accuracy in assigning ages to fur seals. *J. Wildl. Manage.* 34(4):844-52.

Best, P.B. 1969. The sperm whale (*Physeter catodon*) off the west coast of South Africa. 5 Age, growth and mortality. *Investl Rep. Div. Sea Fish. S. Afr.* 79:1-27.

Bjørge, A., Hohn, A.A., Kvam, T., Lockyer, C., Schweder, T. and Aarefjord, H. 1995. Summary Report from the Harbour Porpoise Age Determination Workshop, Oslo, 21-23 May 1990. *Rep. int. Whal. Commn* (special issue) 16:477-96.

Bow, J.M. and Purday, C. 1966. A method of preparing sperm whale teeth for age determination. *Nature, Lond.* 210(5034):437-8.

Brodie, P.F., Geraci, J.W. and St Aubin, D.J. 1990. Dynamics of tooth growth in beluga whales, *Delphinapterus leucas*, and effectiveness of tetracycline as a marker for age determination. pp. 141-8. *In:* T.G. Smith, D.J. St Aubin and J.R. Geraci (eds.) *Advances in Research on the Beluga Whale,* Delphinapterus leucas. Canadian Bulletin of Fisheries and Aquatic Sciences 224, Department of Fisheries and Oceans, Ottawa.

Chang, W.Y.B. 1982. A statistical method for evaluating the reproducibility of age determination. *Can. J. Fish. Aquat. Sci.* 39:1208-0.

Donovan, G., Breiwick, J. and Best, P. 1982. A note on comparative readings of sperm whale teeth. *Rep. int. Whal. Commn* 32:251-2.

Doubleday, W.G. and Bowen, W.D. 1980. Inconsistencies in reading the age of harp seal (*Pagophilus groenlandicus*) teeth, their consequences and a means of reducing biases. NAFO SCR Doc. 80/XI/160. North Atlantic Fisheries Organisation. 24pp.

Evans, K. and Robertson, K. 2001. A note on the preparation of sperm whale (*Physeter macrocephalus*) teeth for age determination. *J. Cetacean Res. Manage.* 3(1):101-7.

Evans, K., Morrice, M., Hindell, M. and Thiele, D. 2002. Three mass strandings of sperm whales, *Physeter macrocephalus*, in southern Australian waters. *Mar. Mammal Sci.* 18(3).

Gambell, R. 1977. Dentinal layer formation in sperm whale teeth. pp. 583-90. *In:* M. Angel (ed.) *A Voyage of Discovery: George Deacon 70th Anniversary Volume*. Pergamon Press Ltd, Oxford.

Hohn, A.A. 1980. Analysis of growth layers in the teeth of *Tursiops truncatus* using light microscopy, microradiography, and SEM. *Rep. int. Whal. Commn* (special issue) 3:155-60.

Hohn, A.A. 1990. Reading between the lines: An analysis of age estimation in dolphins. pp. 575-86. *In:* S. Leatherwood and R.R. Reeves (eds.) *The Bottlenose Dolphin*. Academic Press, San Diego, CA. 653pp.

Hohn, A. and Fernandez, S. 1999. Biases in dolphin age structure due to age estimation technique. *Mar. Mammal Sci.* 15(4):1124-32.

Hohn, A.A., Scott, M.D., Wells, R.S., Sweeney, J.C. and Irvine, A.B. 1989. Growth layers in teeth from known-age, free-ranging bottlenose dolphins. *Mar. Mammal Sci.* 5(4):315-42.

Hui, C.A. 1980. Variability of dentin deposits in *Tursiops truncatus*. *Can. J. Fish. Aquat. Sci.* 37:712-6.

International Whaling Commission. 1967. Report of the Scientific Committee. Appendix E. Sperm Whale Sub-Committee Meeting Report. *Rep. int. Whal. Commn* 17:120-7.

International Whaling Commission. 1969. Report of the Scientific Committee, Appendix H. Report of the Meeting on Age Determination in Whales, Oslo, 1968. *Rep. int. Whal. Commn* 19:131-7.

International Whaling Commission. 1971. Report of the Scientific Committee, Annex D. Report of the Special Meeting on Sperm Whale Biology and Stock Assessments. *Rep. int. Whal. Commn* 21:40-50.

Klevezal, G.A. and Myrick, A.C., Jr. 1984. Marks in tooth dentine of female dolphins (genus *Stenella*) as indicators of parturition. *J. Mammal.* 65(1):103-10.

Lockyer, C. 1980. Age determination studies on *Physeter macrocephalus*. *Rep. int. Whal. Commn* (special issue) 3:216. [Abstract].

Lockyer, C. 1993. A report on patterns of deposition of dentine and cement in teeth of pilot whales, genus *Globicephala*. *Rep. int. Whal. Commn* (special issue) 14:137-61.

Lockyer, C. 1995. A review of factors involved in zonation in odontocete teeth, and an investigation of the likely impact of environmental factors and major life events on harbour porpoise tooth structure. *Rep. int. Whal. Commn* (special issue) 16:511-29.

Mikhalev, Y.A. 1982. Subjective and objective assessments of the lamination in sperm whale teeth. *Rep. int. Whal. Commn* 32:227-33.

Myrick, A.C. 1988. Is tissue resorption and replacement in permanent teeth of mammals caused by stress-induced hypocalcemia? pp. 379-89. *In:* Z. Davidovitch (ed.) *The Biological Mechanisms of Tooth Eruption and Root Resorption*. EBSCO Media, Birmingham, Alabama.

Myrick, A.C., Shallenbeger, E.W., Kang, I. and MacKay, D.B. 1984. Calibration of dental layers in seven captive Hawiian spinner dolphins, *Stenella longirostris*, based on tetracycline labeling. *Fish. Bull.* 82(1):207-25.

Myrick, A.C., Yochem, P.K. and Cornell, L.H. 1988. Toward calibrating dentinal layers in captive killer whales by use of tetracycline labels. *Rit Fisk.* 11:285-96.

Nishiwaki, M., Hibiya, T. and Ohsumi, S.K. 1958. Age study of sperm whale based on reading of tooth laminations. *Sci. Rep. Whales Res. Inst., Tokyo* 13:135-53.

Ohsumi, S. 1971. Some investigations on the school structure of sperm whale. *Sci. Rep. Whales Res. Inst., Tokyo* 23:1-25.

Ohsumi, S. 1977. Age determination and growth of Cetacea. *Honyu-Kagaku* 34(4):54-63.

Ohsumi, S., Kasuya, T. and Nishiwaki, M. 1963. The accumulation rate of dentinal growth layers in the maxilary tooth of the sperm whale. *Sci. Rep. Whales Res. Inst., Tokyo* 16:15-35.

Perrin, W.F. and Myrick, A.C. (eds.). 1980. *Report of the International Whaling Commission (Special Issue 3). Age Determination of Toothed Whales and Sirenians*. International Whaling Commission, Cambridge, UK. 229pp.

Reilly, S.B., Hohn, A.A. and Myrick, A.C. 1983. Precision of age determination of northern offshore spotted dolphins. NOAA Tech. Mem. NOAA-TM-NMFS-SWFSC-35. (Unpublished). iv+27pp. [Available from Southwest Fisheries Science Center, National Marine Fisheries Service, CA].

Rice, D.W., Wolman, A.A., Mate, B.R. and Harvey, J.T. 1986. A mass stranding of sperm whales in Oregon: sex and age composition of the school. *Mar. Mammal Sci.* 2(1):64-9.

Scheffer, V.B. and Myrick, A.C. 1980. A review of studies to 1970 of growth layers in the teeth of marine mammals. *Rep. int. Whal. Commn* (special issue) 3:51-63.

Retinoids in marine mammals and their use as biomarkers of organochlorine compounds

A. BORRELL, V. TORNERO AND A. AGUILAR

Department of Animal Biology and GRUMM, Parc Científic de Barcelona, University of Barcelona, Barcelona, 08028, Spain
Contact e-mail: assump@bio.ub.es

ABSTRACT

Retinoids, also known as vitamin A, are non-endogenous molecules that are essential for a number of physiological processes in mammals. Imbalance of retinoids has been associated with reproductive impairment, embryonic mortality, growth retardation and bone deformities, pathologies in skin and the nervous system, and immune suppression. Mammals cannot produce retinoids so their primary source is dietary. They are absorbed by the small intestine and packaged as retinyl esters in chylomicrons, which enter the circulation and end up mostly in the liver and fatty tissues. Plasma retinoid levels are homeostatically regulated, so they remain constant despite variations in dietary supply or tissue stores. Therefore body depletion of retinoids cannot be reliably assessed through levels in blood, and should be evaluated through concentrations in depot tissues. In marine mammals, the main storage sites for retinoids are liver and blubber. Although not a universal rule, the concentration of retinoids often increases with age in both sexes because of progressive build-up of retinyl esters. In addition, sex often affects retinoid levels, but the nature and magnitude of this effect varies between species and populations. Taxonomic, life-style (particularly dietary) and climatic differences may explain dissimilarities in the effect of age and sex on retinoid levels. For this reason, retinoids can be used to distinguish populations or population components showing distinct dietary, behavioural, or other traits. Disease, particularly when affecting organs of physiological importance or inducing malnutrition, may affect retinoid tissue levels, so care should be taken when studying concentrations in stranded animals. Organochlorine compounds, particularly PCBs, dioxin (TCDDs) and DDTs, increase mobilisation of retinoids from hepatic and extrahepatic storage sites into serum, accompanied by enhanced degradation and elimination of retinoids through urine. In terrestrial mammals, this effect increases retinoid concentration. Conversely, in some species of marine mammals plasma retinoid levels have been reported to decrease when exposure to organochlorines increases, although the physiological mechanisms are unclear. However, given the homeostatic regulation of retinoids in blood, variation in plasma is expected to be less than that in liver or blubber. Because retinoid tissue levels vary in marine mammals even at moderate exposure to organochlorines, and original levels are restored when such exposure decreases or disappears, retinoids may be used as a biomarker of the impact of pollutants on populations. Further research is needed to validate their use, particularly in cetaceans.

KEYWORDS: MARINE MAMMALS; RETINOL; ORGANOCHLORINES; BIOMARKERS

INTRODUCTION

Retinoids, also known as vitamin A, are a family of essential molecules involved in a number of physiological functions in mammals. They are not produced endogenously and can only be acquired from external sources (Blomhoff, 1994); therefore, the capacity of the organism for regulation is limited (Green and Green, 1994). From an environmental perspective, retinoids have attracted attention because some xenobiotic compounds, particularly organochlorine compounds such as PCBs and dioxins, have been found to cause their depletion in mammalian body tissues. This effect may induce severe physiological dysfunction (e.g. Thompson, 1976; Brouwer *et al.*, 1989a; b; Jurek *et al.*, 1990; Håkansson *et al.*, 1991a; b; 1992; Ikegami *et al.*, 1991; Chu *et al.*, 1995; 1996; Kelley *et al.*, 2000). As a result of their physiological role and reactivity to certain chemicals, retinoids have been proposed as a biomarker for exposure to organochlorine and perhaps other pollutants (Peakall, 1992; Jensen *et al.*, 1995; Murk *et al.*, 1998). Given that marine mammals (particularly small predators such as seals, dolphins and porpoises) are subject to extremely high levels of organochlorine exposure, this type of pollution is a potential threat to the conservation of populations (Aguilar *et al.*, 1999; O'Shea and Aguilar, 2001).

Although the liver and other tissues of some large whales have long been recognised as a profitable source of retinoids (see below), information on these compounds in marine mammals is limited. This paper reviews current information on retinoid physiology, natural patterns of variation, and effect of organochlorine exposure in marine mammals. When no information is available, literature on terrestrial mammals is referred to.

CHEMICAL STRUCTURE OF RETINOIDS

Retinoid is a general term referring to a group of closely related compounds whose molecular structure consists of four isoprenoid units joined in a head-to-tail manner. This definition includes compounds such as retinol and retinol derivatives, retinal, retinyl palmitate and retinoic acid.

All-trans-retinol (vitamin A alcohol) is the parent vitamin A compound. It is a fat-soluble primary alcohol of low molecular weight (mw = 286). The aldehyde form, retinal, is found in the retina of the eye, and retinoic acid, a metabolite of vitamin A, is highly active in a number of physiological processes (Wolf, 1984).

PHYSIOLOGY OF RETINOIDS

Although retinoids can be toxic in high concentrations and the adverse effects of hypervitaminosis are documented in both man and animals (Armstrong *et al.*, 1994), in natural conditions the most frequent disorders and pathological effects are produced by low availability of the compound.

Retinoids play an important role in: vision (xerophthalmia and night blindness are both symptoms of its deficiency); the maintenance of the reproductive, endocrine and immune systems; growth and foetal development; and the regulation of the proliferation and differentiation of many cell types. Thus, imbalance of retinoids has been associated with a diversity of anomalies, including reproductive impairment, embryonic mortality, growth retardation and bone deformities, pathologies in skin and the nervous system, and immune suppression (Thompson, 1976; Peakall, 1992). Many of these effects are mediated by the action of retinoic

acid on gene expression (Blomhoff *et al.*, 1991). In addition, retinoids have a protective effect against the development of various cancers (Wolf, 1984).

For mammals, none of which can synthesise retinoids, vitamin A is an essential nutrient. Dietary retinoids are available from two sources: from plants in the form of provitamin A precursor compounds, namely β- (mainly), α- and γ-carotene and cryptoxanthin; and from animal tissues as long-chain retinyl esters.

Once in the digestive system, retinoids are absorbed by the small intestine and packaged as retinyl esters in chylomicrons, which enter the circulation and are taken mostly by the liver. In this organ, chylomicrons are metabolised and retinyl esters are processed for hepatic storage or for secretion as retinol bound to retinol binding protein or RBP (mw = 21000) (Blomhoff, 1994; Green and Green, 1994).

RBP delivers plasma retinoids to target tissues throughout the body (Soprano and Blaner, 1994). Over 95% of RBP-retinol circulates in the blood as a 1:1 molar complex with a second transport protein called transthyretin or TTR (mw = 54980), which also transports thyroid hormones TT4 (Blomhoff, 1994; Green and Green, 1994). It has been established that retinoids recycle among plasma, liver and extrahepatic tissues, since the plasma retinoid turnover is more than one order of magnitude greater than the utilisation rate. The vehicle for retinoid recycling is RBP (Blomhoff *et al.*, 1992; Sommer and West, 1996).

Plasma retinoid levels are constant despite great variation in dietary supply or in liver and extrahepatic tissues stores. Thus, it appears that plasma retinoid levels are homeostatically regulated, ensuring that retinoids are continuously available to vitamin A-dependent cells (Wolf, 1984). As a consequence, body depletion of retinoids cannot be assessed through circulating levels in blood, but should be evaluated through concentrations in depot tissues such as liver and fat. The excretion of retinoids in the urine does not appear to be affected by the retinoid status of the animal itself but by the amount of retinoids available through the diet (Raila *et al.*, 2000).

STORAGE OF RETINOIDS IN TISSUES

The comparative tissue distribution of retinoids in mammals has not been studied systematically. However, surveys available for terrestrial species usually point to the liver as the main storage site, with 50-80% of the body load commonly present in this organ. Extrahepatic tissues such as kidneys, adipose tissue, lung or testis, can also play a significant role in the storage and mobilisation of these compounds (Blaner and Olson, 1994). However, there are dissimilarities among species and/or taxonomic groups. For example, in the Canidae and Mustelidae families, retinoid concentrations in plasma are about 10-50 times higher than in other mammals; indeed, in many mammals such a high level would reflect hypervitaminosis A (Schweigert *et al.*, 1990; 1991b). Kidney retinoid concentration in canids is also high and far exceeds those in the liver; such low hepatic levels would normally be considered an indication of severe vitamin A deficiency in other mammals (Underwood, 1984; Schweigert and Buchholz, 1995). It should be pointed out that the urine of canids contains both retinol and retinyl esters (Schweigert *et al.*, 1991a), while that of human and rats only contains metabolic forms of retinoids, such as retinoic acid (Schweigert and Buchholz, 1995). Therefore, the high level of retinoids observed in the kidney of at least the canids can be associated with this particular form of excretion (Schweigert and Buchholz, 1995). As stated above, generally in terrestrial mammals, the concentration of retinoids in blood is kept constant homeostatically and it decreases only when storage tissues are severely depleted (Wolf, 1984; Blomhoff *et al.*, 1992).

Information on the distribution of retinoids in the body of marine mammals is limited to a few studies that report the concentration in selected tissues from the same individuals (Table 1). There are some data on concentrations in isolated tissues, but these cannot be compared between studies because of substantial variation at individual, population and species levels (see below). The information available suggests that, as is usual in terrestrial mammals, retinoids are extensively stored in the form of retinyl esters in the liver. Indeed, it has long been known that the liver of cetaceans is extremely rich in retinoids (Schmidt-Nielsen *et al.*, 1934), and the interest in obtaining this compound for commercial production of vitamin A led a number of researchers during the first half of the century to investigate its contents in the tissues of large whales (e.g. Klem, 1935; Wetlesen, 1938; Braekkan, 1948; Ishikawa *et al.*, 1948; 1951; Kaneko, 1948; Mori and Saiki, 1950; Tawara and Fukazawa, 1950a; b). A similar richness in hepatic retinoids was later confirmed in pinnipeds (Rodahl and Davies, 1949; Schweigert *et al.*, 1987; Ball *et al.*, 1992; Schweigert and Buchholz, 1995; Käkëla and Hyvärinen, 1997; Käkëla *et al.*, 1997).

However, in marine mammals, blubber is also a significant storage site of retinoids and the concentration of retinoids in the blubber of at least some marine mammals appears to be higher than in comparable fatty tissues of man and other terrestrial mammals (Schweigert *et al.*, 1987). Thermoregulatory and lipid storage needs render fatty tissues of marine mammals to be a substantial proportion of body mass, usually in the range 15-55% and, given the lipophilic nature of retinoids, this allows for massive accumulation of these compounds. The blubber/body mass ratio in marine mammals is inversely scaled, so smaller species tend to have a larger contribution of fatty tissues, and therefore larger relative retinoid stores, than larger species (Ryg *et al.*, 1990; 1993; Aguilar *et al.*, 1999). In grey seals (*Halichoerus grypus*), Schweigert *et al.* (1987) have estimated that blubber accounts for about 40% of total body reserves of retinoids. Borrell *et al.* (1999) found that blubber is also a significant site for retinoid deposition in harbour porpoises (*Phocoena phocoena*) from West Greenland.

Information on retinoid levels in tissues or body organs other than liver and blubber is fragmentary. Mori and Saiki (1950) reported concentrations in the intestine of sperm whales (*Physeter macrocephalus*), Iida *et al.* (1998) in muscle of Antarctic minke whales (*Balaenoptera acutorostrata*), Gregory *et al.* (1955) in the milk of blue whales (*B. musculus*), and Rosas and Lehti (1996) in the milk of Amazon river dolphins (*Inia geoffrensis)*. However, the sample size in these studies was extremely small, often limited to a single individual, and they offer no reliable insight into individual variation. Studies in harp seals (*Pagophilus groenlandicus*), grey seals and common seals (*Phoca vitulina*) indicate that other tissues such as kidneys, lung, retina, pancreas and spleen also have minor shares of the retinoid body content (Rodahl and Davies, 1949).

MAIN FACTORS AFFECTING VARIATION IN TISSUE CONCENTRATIONS

As stated above, retinoids are regulated within individual organisms. However biological traits (e.g. sex, age, diet and body condition, incidence of disease, occurrence of

Table 1

Distribution of retinoids (mean ± SD) in plasma (μg/ml) and other tissues (μg/g tissue) of marine mammals. Only surveys reporting concentrations in more than one tissue have been included (see text).

Species	Location	*n*	Age/Sex (M/F)	Liver	Blubber	Serum	Kidney	Lung	Reference
Harp seal (*Pagophilus groenlandicus*)	Newfoundland	1	Adult	720	3.6	-	1.8	0.9	Rodahl and Davis, 1949
Grey seal (*Halichoerus grypus*)	Pembrokeshire	1	Juvenile	465	1.074	-	4.725	0.75	Rodahl and Davis, 1949
Grey seal (*Halichoerus grypus*)	Sable Island	12	Adult M	502.6 ± 314.9	33.7 ± 10.9	0.26 ± 0.057	-	-	Schweigert *et al.*, 1987
Grey seal (*Halichoerus grypus*)	Sable Island	5	Adult F	264.9 ± 118.4	62.4 ± 3.7	0.41 ± 0.085	-	-	Schweigert *et al.*, 1987
Grey seal (*Halichoerus grypus*)	Sable Island	21	Juvenile	375.7 ± 320.6	21.9 ± 14.8	0.21 ± 0.068	-	-	Schweigert *et al.*, 1987
Grey seal (*Halichoerus grypus*)	Sable Island	6	Adult M	609 ± 395	45 ± 10	0.2 ± 0.1	8 ± 3 (all)	-	Schweigert and Buchholz, 1995
Harbour seal (*Phoca vitulina*)	Wash	1	Juvenile	27	Not detected	-	0.27	0.18	Rodahl and Davis, 1949
Ringed seal (*Pusa hispida*)	Baltic Sea	7-9	Adult	175.3 ± 32.6 (*n*=7)	21.6 ± 3.4 (*n*=9)	-	-	-	Käkelä *et al.*, 1997
Ringed seal (*Pusa hispida*)	Lake Ladoga	4	Juvenile	36.1 ± 7.6	3.1 ± 0.5	-	-	-	Käkelä *et al.*, 1997

lactation) and anthropogenic influences (e.g. environmental pollutants) have a substantial effect on tissue levels and body content of retinoids.

Age

The influence of ageing on retinoids status in terrestrial mammals has been widely studied. Many authors reported an increase in concentrations with age: e.g. liver and blood of rats (Blomhoff *et al.*, 1988); plasma of Florida panthers, *Felis concolor oryi* (Dunbar *et al.*, 1999) and man (Malvy *et al.*, 1993; Stephenson and Gildengorin, 2000), and in the kidney of dogs (Schweigert *et al.*, 1998). However, other surveys revealed either no trend in retinoid levels between age classes, or even decreasing ones. For example, Garry *et al.* (1987) found similar plasma retinoid levels in young and old humans, and Savage *et al.* (1999) reported that age did not affect plasma levels of retinoids in free-ranging African elephants (*Loxodonta africana*). A decrease in serum retinoids was observed by Succari *et al.* (1991) in humans and by Shrestha *et al.* (1998) in female Nepalese elephants (*Elephas maximus*).

Similarly, studies on pinnipeds and cetaceans (Table 2) do not produce consistent results. While many populations showed, both in the liver and in the blubber, an increasing trend in retinoid concentrations with age, others revealed no apparent trend or even a decreasing tendency with age. This variation could not be explained by inter-specific, inter-population or even inter-tissue differences. For example, the studies on ringed seals (*Pusa hispida*) from Lake Saimaa by Käkelä *et al.* (1997) showed a significant positive age-related trend in the blubber and a negative trend in the liver, while those conducted on the same species by the same research group and with an identical sample size ($n = 12$) in Spitsbergen showed the opposite result: a negative trend in the blubber and a positive trend in the liver, although in this case the correlation was non-significant (Table 2).

However, although a general, consistent pattern cannot be deduced from the information available, an increasing trend was the most common finding. This relationship appears to be the result of a decrease in the circulatory clearance of retinoids and other liposoluble compounds with age, coupled with an excess intake of retinoids via diet, which leads to a

Table 2

Age trends in retinoid concentrations observed in tissues of marine mammals. * = only females; ** = significative $p<0.05$; " = statistics not performed; ↑ = positive trend; ↓ = negative trend.

Species	Location	*n*	Liver	Blubber	Reference
Australian fur seal (*Arctocephalus forsteri*)	Australia	24*	↑**		Southcott *et al.* 1974
Grey seals (*Halichoerus grypus*)	Sable Island	65	↑ "	↑ "	Schweigert *et al.* 1987
Hooded seals (*Cystophora cristata*)	Newfoundland	60	↑ "		Rodahl and Davis, 1949
Harp seal (*Pagophilus groenlandicus*)	Newfoundland	145	↑ "		Rodahl and Davis, 1949
Ringed seals (*Pusa hispida*)	Lake Saimaa	12	↓**	↑**	Käkelä *et al.* 1997
Ringed seals (*Pusa hispida*)	Spitsbergen	12	↑	↓	Käkelä *et al.* 1997
Ringed seals (*Pusa hispida*)	Baltic Sea	9	↓	↑	Käkelä *et al.* 1997
Harbour porpoise (*Phocoena phocoena*)	Greenland	100		↑	Borrell *et al.*, 1999

build-up of retinyl ester concentrations with age (Maiani *et al.*, 1989; Krasinski *et al.*, 1990). Although it is not clear why some species or populations do not show this general trend, taxonomic, life-style (particularly dietary) and climatic differences may be responsible.

Sex

Information on sex-related variation in retinoids is even more sparse and less consistent than that for age. In terrestrial mammals, no gender-related differences were observed in circulating concentrations of retinoids in black rhinoceros, *Diceros bicornis* (Ghebremeskel *et al.*, 1988), serum levels in free-ranging African elephants (Savage *et al.*, 1999) or liver and serum concentration in humans (Raica *et al.*, 1972; Succari *et al.*, 1991). Conversely, circulating retinoid levels were reported to be higher in female Florida panthers (Dunbar *et al.*, 1999) but lower in females in some human populations (Krasinski *et al.*, 1989; Stephenson and Gildengorin, 2000). These inter-specific differences may be produced by dissimilarities in types of diet and source of retinoids.

In marine mammals, studies on pinnipeds have often suggested sex-related differences although these varied among tissues and species (Fig. 1). Levels of retinoids were found to be higher in the blubber of adult female grey seals (Schweigert *et al.*, 1987) and in the liver of adult female Australian fur seals (*Arctocephalus forsteri*) (Southcott *et al.*, 1974) than in the corresponding tissues of adult males. However, other surveys have shown the reverse trends. Thus, Rodahl and Davies (1949) found higher concentrations in the liver of male hooded and harp seals than in those of females, and Schweigert *et al.* (1987) found a similar difference in the liver of grey seals. In cetaceans, the only available survey refers to harbour porpoises, in which no significant differences were found between the blubber retinoid concentrations of males and females (Borrell *et al.*, 1999).

It has been suggested that mothers transfer retinoids to their calves during lactation (Simms and Ross, 2000), which would explain the lower levels in the liver of adult females (Schweigert *et al.*, 1987). Milk is a source of essential nutrients, including retinoids. Although studies are limited, marine mammals appear to have relatively higher levels of retinoids in their milk than terrestrial mammals. However, this appears to be due to the high lipid content of the milk in pinnipeds and cetaceans because, when concentrations are expressed as quantity per unit lipid, levels are of the same order of magnitude or even lower than in terrestrial mammals (Schweigert and Stobo, 1994; Debier *et al.*, 1999). Irrespective of this, during lactation, females of both cetaceans and pinnipeds mobilise a large proportion of their blubber reserves, including the blubber-associated retinoid stores. This explains why during lactation, unlike humans, marine mammals may have high levels of circulatory retinoids coupled with lowered stores of retinoids in the blubber and probably other tissues (Schweigert *et al.*, 1987). However, no explanation has been put forward to explain the higher concentrations of males reported in some studies.

Similarly to the age-related variation, it is likely that taxonomic, dietary and life-style dissimilarities between sexes are responsible for sex-related variations. Reproductive activity may be particularly significant in adult individuals because it often involves changes in hormone levels, behavioural traits and diet (see below).

Diet and nutritive condition

Since retinoids are incorporated via food, diet affects tissue levels. However, it is unknown, even in man and laboratory animals, whether body stores of retinoids change as a function of long-term intake of these compounds (Ascherio *et al.*, 1992; Booth *et al.*, 1997; Scrofano *et al.*, 1998). As mentioned above, retinoids in blood are homeostatically controlled when liver stores are sufficient and therefore they only respond to extreme situations, for which reason diet has not been observed to have an effect on them (Blaner and Olson, 1994).

In marine mammals, information on the influence of diet on retinoid status is limited to the study by Käkelä *et al.* (1997), who reported differences in liver and blubber levels between freshwater and marine ringed seals and attributed them to food quality. Differences in diet, as well as climatic or photoperiod dissimilarities may explain variations in retinoid levels between allopatrid populations of the same species. However, such differences may also occur between different components within a single population. For example, variation in diet associated with age, sex or

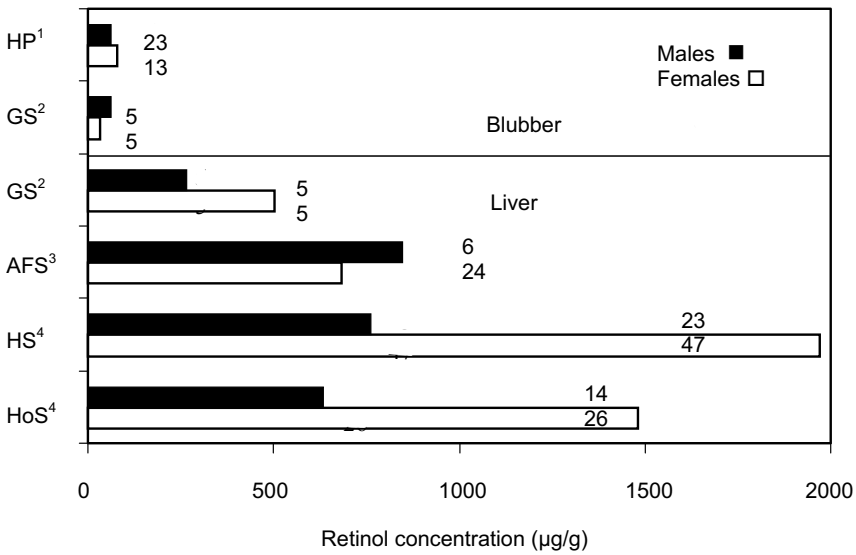


Fig. 1. Sex related variation in retinoid levels (µg/g) in liver and blubber of different marine mammal species. Key: HP = Harbour porpoise; GS = Grey seal; AFS = Australian fur seal; HS = Harp seal; HoS = Hooded seals. References: [1]Borrell *et al.*, 1999; [2]Schweigert *et al.*, 1987; [3]Soutcott *et al.*, 1974; [4]Rodahl and Davies, 1949.

reproductive condition has been reported for many cetaceans and pinnipeds (Seaman *et al.*, 1982; Perez and Mooney, 1986; Stewart and Murie, 1986; Bernard and Hohn, 1989; Recchia and Read, 1989; Rodhouse *et al.*, 1992; Smith and Read, 1992; Clarke *et al.*, 1993). Retinoids thus have the potential to be used to distinguish populations or population components with distinct dietary, behavioural or other traits, provided that the natural sources of variation are properly controlled.

The effect of nutritive condition on retinoid levels is difficult to assess. Neither Rodahl and Davies (1949) or Borrell *et al.* (1999), found any significant effect of condition on the retinoids present in the liver and blubber of hooded and harp seals, or in the blubber of harbour porpoises, respectively. Nevertheless, this conclusion should be treated with caution. The sample examined by Borrell *et al.* (1999) was mainly composed of healthy individuals. While in these conditions, retinoid tissue distribution may remain unaltered. It may require situations of food shortage, massive fat mobilisation (e.g. during migration in large baleen whales or during intensive lactation in some phocids), or starvation caused by disease or other condition, for retinoids to be significantly mobilised, redistributed or excreted. This may be particularly relevant for stranded cetaceans, often found in poor nutritive condition.

The tissue vitamin concentration reflects the essential amounts of these substances necessary for enzymatic and metabolic pathways, coupled with any excess picked up from the environment. The establishment of baseline values of retinoid concentrations is a requisite for the understanding of the chronic effects of toxicity and deficiency (Gelatt *et al.*, 1999).

Disease

Disease, particularly when it affects organs of physiological importance or induces malnutrition, may affect tissue levels of retinoids. However, the information available on this is restricted to humans. Patients suffering from acute or chronic diseases of the liver such as hepatitis, cirrhosis and hepatic tumours have markedly reduced serum levels of RBP, TTR and retinol. Those affected by significant renal disease also show disorders in RBP and retinoid transport, since the kidneys are a major site for RBP catabolism; thus, levels of retinoids increase when excretion is reduced as a consequence of renal tubular damage or reduced glomerular filtration rate of retinol RBP (Goodman, 1984). In addition, sub-normal serum concentrations of RBP and retinoids have been found in patients with a variety of cancers, but it is not clear whether this is a result of protein or energy denutrition (Soprano and Blaner, 1994).

No information is available on disease and retinoids in marine mammals. Given that disease may affect retinoid tissue levels, data from stranded animals in which disease is suspected should not be included in surveys of retinoid status.

EFFECT OF ORGANOCHLORINE POLLUTANTS ON RETINOIDS

Organochlorine compounds can alter retinoid metabolism. However, the biochemical pathway and intensity of the toxic effect appears to vary among species (Håkansson *et al.*, 1991a; Zile, 1992). In general, exposure to PCBs, dioxin (TCDDs) and DDTs leads to depletion of retinoid reserves in mammalian tissue due to increased mobilisation of retinoids from storage sites, especially the liver, and a subsequent increase in their degradation rate (Kelley *et al.*, 2000).

In terrestrial mammals (e.g. rats, otters, minks) feeding on a diet containing toxic organochlorine compounds, the retinol and retinyl ester concentrations in several body organs (liver, depot fat, intestine, lungs and adrenals) have been found to be lower in sample groups exposed to organochlorines than in non-polluted groups. (Brunström *et al.*, 1991; Håkansson *et al.*, 1992; Zile, 1992; Chu *et al.*, 1996; 1998; Murk *et al.*, 1998; Käkëla *et al.*, 1999; Nilsson *et al.*, 2000; Rolland, 2000; Simpson *et al.*, 2000). In contrast, the concentration of retinoids in kidney and, to a lesser extent, in serum, generally increased (Brouwer *et al.*, 1989a; Jurek *et al.*, 1990; Håkansson *et al.*, 1991a; b; Van Birgelen *et al.*, 1994a; b; Chu *et al.*, 1995; Nilsson *et al.*, 2000). This indicates that organochlorines increase mobilisation of retinoids from hepatic and extrahepatic storage sites into serum, accompanied by enhanced degradation and renal elimination of retinoids through urine (Kelley *et al.*, 1998; 2000). Studies on coplanar PCBs and TCDDs have shown that the toxic effect of these compounds is positively correlated with their ability to bind the Ah (arylhydrocarbon) receptor, causing the induction of cytochromes P-450 1A1 and 1A2 (Pelissier *et al.*, 1992; Brouwer, 1995). Thus, it appears that the mixed-function oxidases containing the cytochrome P450s are particularly active in metabolising retinoic acid (Roberts *et al.*, 1979; Ikegami *et al.*, 1991). Moreover, Roberts *et al.* (1992) reported that many rabbit liver cytochrome P-450 isoforms including 2A4, 1A2, 2E1, 2E2, 2C3, 2G1 can catalyse the 4-hydroxylation of both retinol and retinaldehyde. These findings indicate that the decrease in hepatic retinoids storage is related to the induction of cytochrome P-450 and retinoid metabolisation. In laboratory animals exposed to individual PCB congeners, the order of potency in causing reductions in the hepatic contents of retinoids was: PCB 126 > PCB 77 > PCB 153. This order of potency was found to be positively correlated with the ability of each congener to induce cytochrome P450 and with its toxicity measured as weight loss and thymic involution (Chen *et al.*, 1992; Håkansson *et al.*, 1994). In addition, exposure to organochlorines also inhibits the intestinal absorption of ingested vitamin A, thus exacerbating the imbalance produced by the previous effects (Bank *et al.*, 1989).

However, the retinoid depletive effect of these toxic organochlorines can not simply be extrapolated to all organochlorine forms or derivatives. For example, long-term (1 year) experiments conducted with mink fed with methylsulfonyl-PCBs, which are not very AhR-active, did not reveal any effect on retinoid concentrations in tissues (Lund *et al.*, 1999).

Given the evolutionary basis of the physiological processes involved, most of these effects can probably be extended to marine mammals. However, the specific pathways or dynamics may be somewhat different. Thus, most of the studies so far undertaken in three species of pinnipeds and the polar bear (Table 3) have shown a decrease in plasma retinoids when PCB or other organochlorine (OCs) loads increased (Brouwer *et al.*, 1989b; De Swart *et al.*, 1994; Jensen *et al.*, 1995; Beckmen *et al.*, 1997; Skaare *et al.*, 2001). These results originate from studies in both captive and wild populations. In experiments with captive seals, retinoid concentrations returned to normal when animals were fed with slightly contaminated fish (Brouwer *et al.*, 1989b). Unfortunately, only plasma was analysed, so the mechanisms of this decrease were unclear. Given the homeostatic regulation of retinoids in blood, variation in plasma is expected to be lower than in other tissues such as liver or blubber. The only exception (Table 3)

Table 3

Details of studies reporting observed effects of organochlorine pollutants on plasma retinol levels (1) or plasma retinoid levels (2) in marine mammals, including the polar bear.

Species	Location	*n*	Pollutant	Study type	Effect on concentration	References
Harbour seal (*Phoca vitulina*)	-	24	Organochlorines	Experimental	(1) Decrease	Brouwer *et al.*,1989
Harbour seal (*Phoca vitulina*)	-	22	Organochlorines	Experimental	(2) Decrease	De Swart *et al.*, 1994
Northern elephant seal (*Mirounga angustirostris*)	California	31	Organochlorines	Wild	(1) Decrease	Beckmen *et al.*, 1997
Grey seal (pups) (*Halichoerus grypus*)	Norway	51	PCBs	Wild	(2) Decrease	Jenssen *et al.*, 1995
Harbour seal (pups) (*Phoca vitulina*)	British Columbia/ Washington State	61	PCBs	Wild	(1) Decrease (between populations)	Simms *et al.*, 2000
Harbour seal (pups) (*Phoca vitulina*)	British Columbia/ Washington State	37	PCBs	Wild	(1) Increase (in non-nursing pups)	Simms *et al.*, 2000
Polar bear (*Ursus maritimus*)	Svalbard/ Russian Arctic	79	PCBs	Wild	(2) Decrease	Skaare *et al.*, 2001

appears to be the study by Simms *et al.* (2000), which showed that, although retinoid levels in more polluted populations of harbour seal pups were lower than those in a cleaner population, in non-nursing pups levels were positively correlated with organochlorine levels in the blubber. This correlation was explained by the mobilisation of hepatic stores of retinoids into blood and the disruption of the vitamin A transport complex following exposure to milk-derived pollutants, as previously observed in laboratory and terrestrial mammals.

In pinnipeds, hydroxylated PCBs, which are metabolites produced by phase I enzymes, have also been shown to disrupt retinoid transport complexes in plasma, reducing delivery of retinoids to target tissues (Brouwer *et al.*, 1989b; 1998; Ross and Troisi, 2001) as has been seen in terrestrial mammals.

Given that variation in retinoid tissue levels in marine mammals appears to occur even at moderate exposure to organochlorines (Håkansson *et al.*, 1992; Jensen *et al.*, 1995) and that original levels are restored when pollutants disappear or significantly decrease (Brouwer *et al.*, 1989b), retinoids are potentially sensitive biomarkers of organochlorine exposure. However, it is likely that this sensitivity is higher for retinoid reserve tissues, such as blubber, than for blood. In addition, retinoids play a critical role in reproduction and immune competence, two functions through which organochlorines have allegedly impacted marine mammal populations (e.g. see Reijnders *et al.*, 1999). Thus the identification of any potential imbalance of these compounds is relevant to the assessment of the pollutants impact on the involved populations. However, prior to the use of retinoids as biomarkers in ecotoxicological studies, further research is needed to clarify the dynamics of retinoids and their degradation pathways in the tissues of marine mammals, particularly cetaceans.

ACKNOWLEDGEMENTS

This study was funded by the Comisión Interministerial de Ciencia y TecnologíaCICYT (project AMB99/0640), and the Dirección General de Conservación de la NaturalezaDGCONA, Ministerio de Medio Ambiente of Spain. Thanks are due to Joan Canals Boira (Hospital de Bellvitge, Barcelona, Spain), Tinka Murk (Agricultural University, Wageningen, The Netherlands) and an anonymous reviewer for their comments on the manuscript.

REFERENCES

Aguilar, A., Borrell, A. and Pastor, T. 1999. Biological factors affecting variability of persistent pollutant levels in cetaceans. *J. Cetacean Res. Manage.* (special issue) 1:83-116.

Armstrong, R.B., Ashenfelter, K.O., Eckhoff, C., Levin, A.A. and Shapiro, S.S. 1994. General and reproductive toxicology of retinoids. pp. 545-72. *In:* M.B. Sporn, A.B. Roberts and D.S. Goodman (eds.) *The Retinoids: Biology, Chemistry and Medicine.* Raven Press Ltd, New York. 679pp.

Ascherio, A., Stampfer, M.J., Colditz, G.A., Rimm, E.B., Litin, L. and Willett, W.C. 1992. Correlations of vitamin A and E intakes with the plasma concentrations of carotenoids and tocopherols among American men and women. *J. Nutr.* 122(9):1792-801.

Ball, M.D., Nizzi, C.P., Furr, H.C., Olson, J.A. and Oftedal, O.T. 1992. Fatty-acyl esters of retinol (Vitamin A) in the liver of the harp seal (*Phoca groenlandica*), hooded seal (*Cystophora cristata*), and California sea lion (*Zalophus californianus*). *Biochem. Cell Biol.* 70:809-13.

Bank, P.A., Cullum, M.E., Jensen, R.K. and Zile, M.H. 1989. Effect of hexachlorobiphenyl on vitamin A homeostasis in the rat. *Biochem. Biophys. Acta* 990:306-14.

Beckmen, K.B., Lowenstine, L.J., Newman, J., Hill, J., Hanni, K. and Gerber, J. 1997. Clinical and pathological characterisation of northern elephant seal skin disease. *J. Wildl. Dis.* 33:438-49.

Bernard, H.J. and Hohn, A.A. 1989. Differences in feeding habits between pregnant and lactating spotted dolphins (*Stenella attenuata*). *J. Mammal.* 70(1):211-5.

Blaner, W.S. and Olson, J.A. 1994. Retinol and retinoic acid metabolism. pp. 229-55. *In:* M.B. Sporn, A.B. Roberts and D.S. Goodman (eds.) *The Retinoids: Biology, Chemistry and Medicine*. Raven Press Ltd, New York. 679pp.

Blomhoff, R. 1994. Overview of vitamin A metabolism and function. pp. 1-35. *In:* R. Blomhoff (ed.) *Vitamin A in Health and Disease*. Marcel Dekker, New York. 677pp.

Blomhoff, R., Berg, T. and Norum, K.R. 1988. Distribution of retinol in rat liver cells: effect of age, sex and nutritional status. *Br. J. Nutr.* 60:233-9.

Blomhoff, R., Green, M.H., Green, J.B., Berg, T. and Norum, K.R. 1991. Vitamin A metabolism: new perspectives on absorption, transport and storage. *Physiol. Rev.* 71(4):951-90.

Blomhoff, R., Green, M.H. and Norum, K.R. 1992. Vitamin A: Physiological and biochemical processing. *Annu. Rev. Nutr.* 12:37-57.

Booth, S.L., Tucker, K.L., McKeown, N.M., Davidson, K.W., Dallaal, G.E. and Sadowski, J.A. 1997. Relationships between dietary intakes and fasting plasma concentrations of fat-soluble vitamins in humans. *J. Nutr.* 127(4):587-92.

Borrell, A., Cantos, G., Aguilar, A., Lockyer, C., Brouwer, A., Heide-Jørgensen, M.P., Jensen, J. and Spenkelink, B. 1999. Patterns of variability of retinol levels in a harbour porpoise population from an unpolluted environment. *Mar. Ecol. Prog. Ser.* 185:85-92.

Braekkan, O.R. 1948. Vitamins in whale liver. *Sci. Rep. Mar. Biol. Res.* 32:1-25.

Brouwer, A. 1995. Metabolism of xenobiotics in laboratory animals and widllife species: potential impact on physiology and health. Paper SC/M95/P8 presented to the IWC Scientific Committee Workshop on Chemical Pollutants and Cetaceans, March 1995,

Bergen (unpublished). [Paper available from the Office of this Journal].

Brouwer, A., Håkansson, H., Kukler, A., Van den Berg, K.J. and Ahlborg, U.G. 1989a. Marked alterations in retinoid homeostasis of Sprague-Dawley rats induced by a single i.p dose of 10μg/kg of 2,3,7,8-tetrachlorodibenzo-p-dioxin. *Toxicology* 58:267-83.

Brouwer, A., Reijnders, P.J.H. and Koeman, J.H. 1989b. Polychlorinated biphenyl (PCB)-contaminated fish induces vitamin A and thyroid hormone deficiency in the common seal (*Phoca vitulina*). *Aquat. Toxicol.* 15:99-106.

Brouwer, A., Morse, D.C., Lans, M.C., Schuur, A.G., Murk, A.J., Wehler, E.K., Bergman, A. and Visser, T.J. 1998. Interactions of persistent environmental organohalogens with the thyroid hormone system: mechanisms and possible consequences for animal and human health. *Toxicol. Ind. Health* 14:59-84.

Brunström, B., Håkansson, H. and Lundberg, K. 1991. Effects of a technical PCB preparation and fractions thereof on ethoxyresorufin *O*-deethylase activity, vitamin A levels and thymic development in the mink (*Mustela vision*). *Pharmacol. Toxicol.* 69:421-6.

Chen, L.-C., Berberian, I., Koch, B., Mercier, M., Azais-Braesco, V., Glauert, H.P., Chow, C.K. and Robertson, L.W. 1992. Polychlorinated and polybrominated biphenyl congeners and retinoid levels in rat tissues: structure-activity relationships. *Toxicol. Appl. Pharmacol.* 114:47-55.

Chu, I., Villeneuve, D.C., Yagminas, A., Lecavalier, P., Håkansson, H., Ahlborg, U.G., Valli, V.E., Kennedy, S.W., Bergman, A., Seegal, R.F. and Feeley, M. 1995. Toxicity of PCB 77 (3,3',4,4'-tetrachlorobiphenyl) and PCB 118 (2,3',4,4',5-pentaclorobiphenyl) in the rat following subchronic dietary exposure. *Fundam. Appl. Toxicol.* 26(2):282-92.

Chu, I., Villeneuve, D.C., Yagminas, A., Lecavalier, P., Poon, R., Feeley, M., Kennedy, S.W., Seegal, R.F., Håkansson, H., Ahlborg, U.G., Valli, V.E. and Bergman, A. 1996. Toxicity of 2,2',4,4',5,5'-hexachlorobiphenyl in rats: Effects following 90-day oral exposure. *J. Appl. Toxicol.* 16(2):121-8.

Chu, I., Poon, R., Yagminas, A., Lecavalier, P., Håkansson, H., Valli, V.E., Kennedy, S.W., Bergman, A., Seegal, R.F. and Feeley, M. 1998. Subchronic toxicity of PCB 105 (2,3,3,4,4',-pentachlorobiphenyl) in rats. *J. Appl. Toxicol.* 18(4):285-92.

Clarke, M.R., Martins, H.R. and Pascoe, P. 1993. The diet of sperm whales (*Physeter macrocephalus* Linnaeus 1758) off the Azores. *Philos. Trans. R. Soc. Lond. B. (Biol. Sci.)* 339(1287):67-82.

Debier, C., Kovacs, K.M., Lydersen, C., Mignolet, E. and Larondelle, Y. 1999. Vitamin E and vitamin A contents, fatty acid profiles, and gross composition of harp and hooded seal milk through lactation. *Can. J. Zool.* 77:952-8.

De Swart, R.L., Ross, P.S., Vedder, L.J., Timmerman, H.H., Heisterkamp, S.H., Van Loveren, H., Vos, J.G., Reijnders, P.J.H. and Osterhaus, A.D.M.E. 1994. Impairment of immune function in harbour seals (*Phoca vitulina*) feeding on fish from polluted waters. *Ambio* 23:155-9.

Dunbar, M.R., Cunningham, M.W. and Linda, S.B. 1999. Vitamin A concentrations in serum and liver from Florida panthers. *J. Wildl. Dis.* 35(2):171-7.

Garry, P.J., Hunt, W.C., Bandrofchak, J.L., Vanderjagt, D. and Goodwin, J.S. 1987. Vitamin A intake and plasma retinol levels in healthy elderly men and women. *Am. J. Clin. Nutr.* 46:989-94.

Gelatt, T.S., Arendt, T., Murphy, M.S. and Siniff, D.B. 1999. Baseline levels of selected minerals and fat-soluble vitamins in Weddel seals (*Leptonychotes weddelli*) from Erebus Bay, McMurdo Sound, Antarctica. *Mar. Poll. Bull.* 38:1251-8.

Ghebremeskel, K., Williams, G., Lewist, J.C.M. and Du Toit, M. 1988. Serum alpha tocopherol, all-trans retinol, total lipids and cholesterol in the black rhinoceros (*Diceros bicornis*). *Comp. Biochem. Physiol. A* 91(2):343-6.

Goodman, D.S. 1984. Plasma retinol-binding protein. pp. 44-88. *In:* M.B. Sporn, A.B. Roberts and D.S. Goodman (eds.) Vol. 2. *The Retinoids*. Academic Press, Orlando, FL.

Green, M.H. and Green, J.B. 1994. Dynamics and control of plasma retinol. pp. 119-33. *In:* R. Blomhoff (ed.) *Vitamin A in Health and Disease*. Marcel Dekker, New York. 677pp.

Gregory, M.E., Kon, S.K., Rowland, S.J. and Thompson, S.Y. 1955. The composition of the milk of the blue whale. *J. Dairy Res.* 22:108-12.

Håkansson, H., Johansson, L., Manzoor, E. and Ahlborg, U.G. 1991a. Effects of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) on the vitamin A status of Hartley guinea pigs, Sprague-Dawley rats, C57B1/6 mice, DBA/2 mice and Golden Syrian hamsters. *J. Nutr. Sci. Vitaminol.* 37(2):117-38.

Håkansson, H., Manzoor, E. and Ahlborg, U.G. 1991b. Interaction between dietary vitamin A and single oral doses of 2,3,7,8-tetrachlorodebenzo-p-dioxin (TCDD) on the TCDD-induced toxicity and on the vitamin A status in the rat. *J. Nutr. Sci. Vitaminol.* 37(3):239-56.

Håkansson, H., Manzoor, E. and Ahlborg, U.G. 1992. Effects of technical PCB preparations and fractions thereof on vitamin A levels in the mink (*Mustela vison*). *Ambio* 21(8):588-90.

Håkansson, H., Manzoor, E., Trossvik, C., Ahlborg, U.G., Chu, I. and Villeneuve, D. 1994. Effect on tissue vitamin A levels in the rat following subchronic exposure to four individual PCB congeners (IUPAC 77, 118, 126 and 153). *Chemosphere* 29:2309-13.

Iida, H., Murata, Y., Matsumoto, G., Toda, S., Yamashita, Y. and Yokoyama, M. 1998. Chemical composition of the edible parts of minke whale *Balaenoptera acutorostrata*. *Bull. Nat. Res. Inst. Fish. Sci.* 11:27-36.

Ikegami, S., Tsuchihashi, F. and Nishide, E. 1991. Decrease in hepatic storage of vitamin A and induction of cytochrome P-450 by dietary organochlorine pesticides in rats. *J. Food Hyg. Soc. Japan* 32(1):1-7.

Ishikawa, S., Omote, Y. and Soma, Y. 1948. Analytical distillation of vitamin A in the whale liver oil. *Sci. Rep. Whales Res. Inst., Tokyo* 2:35-41.

Ishikawa, S., Omote, Y. and Okuda, H. 1951. Substances related to vitamin A in the whale liver oil. *Sci. Rep. Whales Res. Inst., Tokyo* 5:53-9.

Jenssen, B.M., Skaare, J.U., Woldstad, S., Nastad, A.T., Haugen, O., Kloven, B. and Sormo, E.G. 1995. Biomarkers in blood to assess effects of polychlorinated biphenyls in free-living grey seal pups. pp. 607-15. *In:* Blix (ed.) *Whales, Seals, Fish and Man: Proceedings of the International symposium on the Biology of Marine Mammals in the Northeast Atlantic. Tromso, Norway, 29 November –1 December 1994*.

Jurek, M.A., Powers, R.H., Gilbert, L.G. and Aust, S.D. 1990. The effect of TCDD on acyl coenzyme A:retinol acyltransferase activity and vitamin A accumulation in the kidney of male Sprague-Dawley rats. *J. Biochem. Toxicol.* 5(3):155-60.

Käkëla, R. and Hyvärinen, H. 1997. Fatty acid compositions and vitamin A and E status of endangered freshwater ringed seals (from Finland and Western Russia) differ from those of marine seals. *Rec. Res. Devel. Lipids Res.* 1:1-7.

Käkëla, R., Hyvärinen, H. and Käkëla, A. 1997. Vitamins A-1 (retinol), A-2 (3,4 didehydroretinol) and E (alpha-tocopherol) in the liver and blubber of lacustrine and marine ringed seals (*Phoca hispida* sp.). *Comp. Biochem. Physiol. B* 116:27-33.

Käkëla, R., Käkëla, A., Hyvärinen, H., Asikainen, J. and Dahl, S.K. 1999. Vitamins A_1, A_2 and E in minks exposed to polychlorinated biphenyls (*Aroclor 1242*) and copper, via diet based on freshwater or marine fish. *Environ. Toxicol. Chem.* 18:2595-9.

Kaneko, A. 1948. Molecular distillation of fin whale liver oil. *Sci. Rep. Whales Res. Inst., Tokyo* 2:46-50.

Kelley, S.K., Milsson, C.B., Green, M.H., Green, J.B. and Håkansson, H. 1998. Use of model-based compartmental analysis to study effects of 2,3,7,8-tetrachlorodibenzo-p-dioxin on vitamin A kinetics in rats. *Toxicol. Sci.* 44(1):1-13.

Kelley, S.K., Nilsson, C.B., Green, M.H., Green, J.B. and Håkansson, H. 2000. Mobilisation of vitamin A stores in rats after administration of 2,3,7,8-tetrachlorodibenzo-p-dioxin: A kinetic analysis. *Toxicol. Sci.* 55(2):478-84.

Klem, A. 1935. Studies on the biochemistry of whale oils. *Hvalråd. Skr.* 11:49-108.

Krasinski, S.D., Russell, R.M. and Otradovec, C.L. 1989. Relationship of vitamin A and vitamin E intake to fasting plasma retinol, retinol binding protein, retinyl esters, carotene, alpha-tocopherol and cholesterol among elderly people and young adults: increased plasma retinyl esters among vitamin A supplement users. *Am. J. Clin. Nutr.* 49:112-20.

Krasinski, S.D., Cohn, J.S., Shaefer, E.J. and Russel, R.M. 1990. Postprandial plasma retinyl ester response is greater in older subjects compared with younger subjects: evidence for delayed plasma clearance of intestinal lipoproteins. *Clin. Invest.* 85:883-92.

Lund, B.O., Örberg, J., Bergman, Å., Larsson, C., Bergman, A., Bäcklin, B.M., Håkansson, H., Madej, A., Brouwer, A. and Brunström, B. 1999. Chronic and reproductive toxicity of a mixture of 15 methylsulfonyl-polychlorinated biphenyls and 3-methylsulfonyl-2,2-bis-(4-chlorophyenyl)-1,1-dichloroethene in mink (*Mustela vison*). *Environ. Toxicol. Chem.* 18:292-8.

Maiani, G., Mobarhan, S., Ceccanti, M., Ranaldi, L., Gettner, S. and Bowen, P. 1989. Beta-carotene serum responses in younger and elderly females. *Eur. J. Clin. Nutr.* 43:749-61.

Malvy, D.J.M., Buurtschy, B., Dostalova, L. and Amedee-Manesme, O. 1993. Serum retinol, beta-carotene, alpha-tocopherol and cholesterol in healthy French children. *Int. J. Epidemiol.* 22(2):237-46.

Mori, T. and Saiki, M. 1950. Properties of fats and oils contained in various parts of a sperm whale body. *Sci. Rep. Whales Res. Inst., Tokyo* 3:79-84.

Murk, A.J., Leonards, P.E.G., Van Hattum, B., Luit, R., Van de Weiden, M.E.J. and Smit, M. 1998. Application of biomarkers from exposure and effect of polyhalogenated aromatic hydrocarbons in naturally exposed European otters (*Lutra lutra*). *Environ. Toxicol. Pharmacol.* 6:91-102.

Nilsson, C.B., Hoegberg, P., Trossvik, C., Azais-Braesco, V., Blaner, W.S., Fex, G., Harrison, E.H., Nau, H., Schmidt, C.K., Van Bennekum, A.M. and Håkansson, H. 2000. 2,3,7,8-tetrachrorodibenzo-p-dioxin increases serum and kidney retinoic acid levels and kidney retinol esterification in the rat. *Toxicol. Appl. Pharmacol.* 169(2):121-31.

O'Shea, T.J. and Aguilar, A. 2001. Cetacea and Sirenia. pp. 427-96. *In:* R.F. Shore and B.A. Rattner (eds.) *Ecotoxicology of Wild Mammals*. John Wiley and Sons Ltd, New York, NY. 730pp.

Peakall, D.B. 1992. *Animal Biomarkers as Pollution Indicators*. Chapman and Hall, London. 291pp.

Pelissier, M.A., Siess, M.H., Lhuissier, M., Grolier, P., Suschetet, M., Narbonnes, J.F., Albrecht, R. and Robertson, L.W. 1992. Effect of prototypic polychlorinated biphenyls on hepatic and renal vitamin contents and drug-metabolising enzymes in rats fed diets containing low or high levels of retinyl palmitate. *Chem. Res. Toxicol.* 30(8):723-9.

Perez, M.A. and Mooney, E.E. 1986. Increased food and energy consumption of lactating northern fur seals, *Callorhinus ursinus*. *Fish. Bull.* 84(2):371-81.

Raica, N., Scott, J., Lowry, L. and Sauberlich, H.E. 1972. Vitamin A concentration in human tissues collected from five areas in the United States. *Am. J. Clin. Nutr.* 25:291-6.

Raila, J., Bucholz, I., Auerle, H., Raila, G., Schoon, H.A. and Schwigert, F.J. 2000. The distribution of vitamin A and retinol-binding protein in the blood plasma, urine, liver and kidneys of carnivores. *Vet. Res.* 31:541-51.

Recchia, C.A. and Read, A.J. 1989. Stomach contents of harbour porpoises, *Phocoena phocoena* (L.), from the Bay of Fundy. *Can. J. Zool.* 67:2140-6.

Reijnders, P.J.H., Donovan, G.P., Aguilar, A. and BjØrge, A. 1999. Report of the Workshop on Chemical Pollution and Cetaceans, March 1995, Bergen, Norway. *J. Cetacean Res. Manage.* (special issue) 1:1-42.

Roberts, A.B., Nichols, M.D., Newton, D.L. and Sporn, M.B. 1979. In vitro metabolism of retinoic acid in hamster intestine and liver. *J. Biol. Chem.* 254:6296-302.

Roberts, E.S., Vaz, A.D.N. and Coon, M.J. 1992. Role of isozymes of rabbit microsomal cytochrome p450 in the metabolism of retinoic acid, retinol and retinal. *Mol. Pharmacol.* 41:427-33.

Rodahl, K. and Davies, A.W. 1949. Vitamin A in seals. *Biochemistry* 45:408-12.

Rodhouse, P.G., Arnbom, T.R., Fedak, M.A., Yeatman, J. and Murray, W.A. 1992. Cephalopod prey of the southern elephant seal, *Mirounga leonina* L. *Can. J. Zool.* 70:1007-15.

Rolland, R. 2000. A review of chemically-induced alterations in thyroid and vitamin A status from field studies of wildlife and fish. *J. Wildl. Dis.* 36(4):615-35.

Rosas, F.C.W. and Lethi, K.K. 1996. Nutritional and mercury content of milk of the Amazon river dolphin, *Inia geoffrensis*. *Comp. Biochem. Physiol. A* 115(2):117-9.

Ross, P.S. and Troisi, G.M. 2001. Pinnipedia. pp. 371-426. *In:* R.F. Shore and B.A. Ratner (eds.) *Ecotoxicology of Wild Mammals*. John Wiley and Sons Ltd, New York, NY. 730pp.

Ryg, M., Lydersen, C., Markussen, N.H., Smith, T.G. and Oritsland, N.A. 1990. Estimating the blubber content of phocid seals. *Can. J. Fish. Aquat. Sci.* 47(6):1223-7.

Ryg, M., Lydersen, C., Knutsen, L.O., Bjorge, A., Smith, T.G. and Oritsland, N.A. 1993. Scaling of insulation in seals and whales. *J. Zool., Lond.* 230:193-206.

Savage, A., Leonnng, K.M., Grobler, D., Lehhnnhardt, J., Dierenfeld, E.S., Stevens, E.F. and Aebischer, C. 1999. Circulating levels of alpha-tocopherol and retinol in free-ranging African elephants (*Loxodonta africana*). *Zoo Biol.* 18:319-23.

Schmidt-Nielsen, S., Flood, A. and Stene, J. 1934. Ueber Grösse und Vitamingehalt der Leber verschiedener Tiere. *K. Nor. Vidensk. Selsk. For.* 7:81. [In Norwegian].

Schweigert, F.J. and Buchholz, Y. 1995. Vitamin A metabolism in carnivores with special reference to fur bearing animals. *Scientifur* 19(4):305-7.

Schweigert, F.J. and Stobo, W.T. 1994. Transfer of fat-soluble vitamins and PCBs from mother to pups in grey seals (*Halichoerus grypus*). *Comp. Biochem. Physiol.* 109:11-117.

Schweigert, F.J., Stobo, W.T. and Zucker, H. 1987. Vitamin A status in the grey seal (*Halichoerus grypus*) on Sable Island. *Int. J. Vitam. Nutr. Res.* 57:239-45.

Schweigert, F.J., Ryder, O.A., Rambeck, W.A. and Zucker, H. 1990. The majority of vitamin A is transported as retinyl esters in the blood of most carnivores. *Comp. Biochem. Physiol. B* 95A(4):573-8.

Schweigert, F.J., Thomann, E. and Zucker, H. 1991a. Vitamin A in the urine of carnivores. *Int. J. Vitam. Nutr. Res.* 61:110-3.

Schweigert, F.J., Uehlein-Harrel, S., Hegel, G.V. and Wiesner, H. 1991b. Vitamin A (retinol and retinyl esters), alpha-tocopherol and lipid levels in plasma of captive wild mammals and birds. *J. Vet. Med. Sci.* 38:35-42.

Schweigert, F.J., Bucholz, I. and Bonitz, K. 1998. Effect of age on the levels of retinol and retinyl esters in blood plasma, liver and kidney of dogs. *Int. J. Vitam. Nutr. Res.* 68:237-41.

Scrofano, M.M., Jhngen-Hodge, J., Nowell, T.R., Jr., Gong, X., Smith, D.E., Perrone, G., Asmmundsson, G., Dallal, G., Gindleesky, B., Mura, C.V. and Taylor, A. 1998. The effects of ageing and calorie restriction on plasma nutrient levels in male and female Emory mice. *Mech. Ageing Develop.* 105:31-44.

Seaman, G.A., Lowry, L.F. and Frost, K.J. 1982. Foods of belukha whales *Delphinapterus leucas* in Western Alaska. *Cetology* 44:1-19.

Shrestha, S.P., Ullrey, D.E., Bernard, J.B., Wemmer, C. and Kraemer, D.C. 1998. Plasma vitamin E and other analyte levels in Nepalese camp elephants (*Elephas maximus*). *J. Zoo Wildl. Med.* 29(3):269-78.

Simms, W. and Ross, P.S. 2000. Developmental changes in circulatory vitamin A (retinol) and its transport proteins in free-ranging harbour seal (*Phoca vitulina*) pups. *Can. J. Zool.* 78:1862-8.

Simms, W., Jeffries, S.J., Ikonomou, M.G. and Ross, P.S. 2000. Contaminant-related disruption of vitamin A dynamics in free-ranging harbor seal (*Phoca vitulina*) pups from British Columbia, Canada and Washington State, USA. *Environ. Toxicol. Chem.* 19:2844-9.

Simpson, V.R., Bain, M.S., Brown, R., Brown, B.F. and Lacey, R.F. 2000. A long-term study of vitamin A and polychlorinated hydrocarbon levels in otters (*Lutra lutra*) in south west England. *Environ. Pollut.* 110:267-75.

Skaare, J.U., Bernhoft, A., Wiig, O., Norum, K.R., Haug, E., Eide, D.M. and Derocher, A.E. 2001. Relationships between plasma levels or organochlorines, retinol and thyroid hormones from polar bears (*Ursus maritimus*) at Svalbard. *J. Toxicol. Environ. Health* 62(4):227-41.

Smith, R.J. and Read, A.J. 1992. Consumption of euphausiids by harbour porpoise (*Phocoena phocoena*) calves in the Bay of Fundy. *Can. J. Zool.* 70(8):1629-32.

Sommer, A. and West, K.P., Jr. 1996. *Vitamin A Deficiency: Health, Survivial and Vision*. Oxford University Press Inc, New York, NY. 242pp.

Soprano, D.R. and Blaner, W.S. 1994. Plasma retinol binding protein. pp. 257-81. *In:* M.B. Sporn, A.B. Roberts and D.S. Goodman (eds.) *The Retinoids: Biology, Chemistry and Medicine*. Raven Press, New York. 679pp.

Southcott, R.V., Chesterfield, N.J. and Warneke, R.M. 1974. The vitamin A content of the liver of the Australian fur seal, *Arctocephalus pusillus doriferus*. *Aust. Wildl. Res.* 1:145-9.

Stephenson, C.B. and Gildengorin, G. 2000. Serum retinol, the acute phase reponse, and the apparent misclassification of vitamin A status in the third National Health and Nutrition Examination Survey. *Am. J. Clin. Nutr.* 72(5):1170-8.

Stewart, R.E.A. and Murie, D.J. 1986. Food habits of lactating harp seals (*Phoca groenlandica*) in the Gulf of St. Lawrence in March. *J. Mammal.* 67(1):186-8.

Succari, M., Garric, B., Ponteziere, C., Miocque, M. and Cals, M.J. 1991. Influence of sex and age on vitamin A and E status. *Age and Ageing* 20:413-6.

Tawara, T. and Fukazawa, R. 1950a. Studies in kitol. II. Influence of kitol fraction on the determination of the International Unit of Vitamin A. *Sci. Rep. Whales Res. Inst., Tokyo* 3:89-91.

Tawara, T. and Fukazawa, R. 1950b. Studies on kitol. I. Preparation of kitol from whale liver oil. *Sci. Rep. Whales Res. Inst., Tokyo* 3:85-8.

Thompson, J.N. 1976. Fat-soluble vitamins. *Comp. Anim. Nutr.* 1:99-135.

Underwood, B.A. 1984. Vitamin A in animal and human nutrition. pp. 282-392. *In:* M.B. Sporn, A.B. Roberts and D.S. Goodman (eds.) Vol. 2. *The Retinoids*. Academic Press, Orlando, FL.

Van Birgelen, A.P.J.M., Van Der Kolk, J., Fase, K.M., Bol, I., Poiger, H., Brouwer, A. and Van Den Berg, M. 1994a. Toxic potency of 3,3',4,4',5-pentachlorobiphenyl relative to and in combination with 2,3,7,8-tetrachlordibenzo-p-dioxin in a subchronic feeding study in the rat. *Toxicol. Appl. Pharmacol.* 127:209-21.

Van Birgelen, A.P.J.M., Van Der Kolk, J., Fase, K.M., Bol, I., Poiger, H., Van Den Berg, M. and Brouwer, A. 1994b. Toxic potency of 2,3,3',4,4',5-hexachlorobiphenyl relative to and in combination with 2,3,7,8-tetrachlorodibenzo-p-dioxin in a subchronic feeding study in the rat. *Toxicol. Appl. Pharmacol.* 126:202-13.

Wetlesen, C.U. 1938. Whale liver. *Norsk Hvalfangsttid.* 27:262-4.

Wolf, G. 1984. Multiple functions of vitamin A. *Physiol. Rev.* 64:873-973.

Zile, M.H. 1992. Vitamin A homeostasis endangered by environmental pollutants. *Proceed. Soc. Exper. Biol. Med.* 201:141-53.

A note on a computer-based system for theodolite tracking of cetaceans

GLENN GAILEY and JOEL G. ORTEGA-ORTIZ

Marine Mammal Research Program, Texas A&M University, 4700 Avenue U. Building 303, Galveston, TX 77551 USA
Contact e-mail: Gaileyg@tamug.tamu.edu

ABSTRACT

Theodolites represent a non-invasive shore-based tool for obtaining data on cetacean movement patterns, habitat use and behavioural disturbance. Despite the common use of theodolites as research tools, relatively few computer-based systems exist to assist researchers with collection of theodolite derived data and the analysis of such information. A recently developed computer program named 'Pythagoras', provides an efficient and user-friendly tool for collecting, managing and subsequent analysis of data obtained with theodolites. *Pythagoras* provides location of user-defined fix types (e.g. whales, dolphins, boats, etc.) and has a dynamic interface, that can be customised to fit site-specific research needs. Additional information (behaviour, group size and environmental conditions) can be stored with each theodolite fix. Tracking data are immediately available in the form of a real-time graphic representation. All collected data are stored in *Microsoft Access* and can be exported as *Microsoft Excel*, *ArcInfo*, *Surfer*, *MATLAB*, or delimited text file formats. An analysis module is included to calculate linearity, reorientation rate and leg speed for each track, and distance and orientation between two or more tracklines. Behavioural data are analysed for frequency, time intervals (i.e. blow interval), duration (i.e. surface time) and rate (number per minute) of particular behaviours. Several other computer-based theodolite systems are reviewed here to evaluate their potential benefits and limitations as a means of providing a basis for future developments.

KEYWORDS: TECHNIQUE; MOVEMENTS; DISTRIBUTION; MANAGEMENT; SURVEY–SHORE-BASED; BEHAVIOUR

INTRODUCTION

A theodolite is a surveyor's instrument that measures horizontal and vertical angles with high precision and can be successfully used to determine the distance from the observation site to an object. When placed on an elevated shore-based vantage point, theodolites have been used as a research tool to obtain data on cetacean movement patterns, behaviour, distribution and habitat use (e.g. Würsig, 1978; Bejder, 1997; Ward, 1999; Yin, 1999; Brown, 2000). Theodolite derived data are collected in a completely non-invasive manner, as described by Würsig *et al.* (1991). Both cetaceans and other objects, for instance boats, can be tracked, and interactions between them can be continually and accurately monitored. This aspect makes a theodolite a useful tool for conservation and management research, especially in regard to monitoring potential human-related impacts on marine mammals.

Although this technique has been used for almost 30 years and despite the increase in digital theodolite use for cetacean studies, relatively few computer-based theodolite programs exist to assist researchers in collecting, managing and analysing theodolite data. One of the first theodolite programs for cetacean research was developed by Wolitzky to assist Würsig (1978) in analysing the associated data on dusky dolphins (*Lagenorhynchus obscurus*) off Argentina. Cipriano (1990) created a program called 'T-Trak' as a tool to help analyse data collected with a theodolite for dusky dolphins off New Zealand. In the early 1990s, a Macintosh program named 'Aardvark' was developed by Harold Mills to study humpback whales (*Megaptera novaeangliae*) off Hawaii. A recent program called 'Cyclops' was also created to study humpback whales off Australia (Kniest *et al.*, 2000).

A computer-based system benefits theodolite studies in many ways. For instance, angles can be recorded accurately and efficiently. In addition, real-time calculations of distance to the object and its geographic location can be performed, and trackline(s) can be visually displayed, allowing for rapid corrections of possible tracking errors. Moreover, once data are collected, a computer-based system reduces the time needed for further management and analysis.

This paper describes the fundamental components of a recent computer-based theodolite program, *Pythagoras*, and compares this system to other available theodolite programs.

DESCRIPTION

General

The program *Pythagoras* was designed to communicate with a digital theodolite and provide a dynamic and user-friendly interface. The system collects, manages and analyses theodolite data and calculates distance, bearing and geographic location information in real-time (Gailey and Ortega-Ortiz, 2000). *Pythagoras* was written in *Microsoft Visual Basic* with *Microsoft Access* as the database structure and management component of the program. The program stores specified theodolite station information, such as the observation platform height, geographical position and reference azimuth for multiple theodolite stations. A dynamic interface allows users to define 'fix type' objects, such as dolphins, whales and boats, and the behaviours associated with each defined object as well as environmental and other data not related to the fix itself (i.e. group size). Options for focal group or individual behavioural data collection (Martin and Bateson, 1993) are incorporated to provide detailed records of behavioural events of object(s) being tracked. Tide height can affect the accuracy of the distance estimations, and therefore is an important environmental variable to be considered. Tide height data can be imported *a priori* or *post hoc* to utilise predicted or real tide height values, respectively. Geographic information on the researcher's study area can be imported with

geographic information system (GIS) digital vector line maps, such as *ArcInfo* ungenerated format, *Surfer*, *MapInfo* and *MATLAB*.

Data collection

Pythagoras uses the connected theodolite's horizontal and vertical angle readings, and the selected object type to perform calculations of distance, bearing and location upon each newly recorded entry. The recording of the theodolite's angles is subsequently referred to as a 'fix'. The distance calculation performed for each fixed object incorporates the station's geographic position (latitude, longitude), theodolite angle readings, observer's height above sea level and tide height. A modified version of the distance approximation proposed by Lerczak and Hobbs (1998) was used to calculate sighting distances from angular readings (in radians) of shore-based marine mammal surveys, which corrects for the curvature of the earth (Fig. 1).

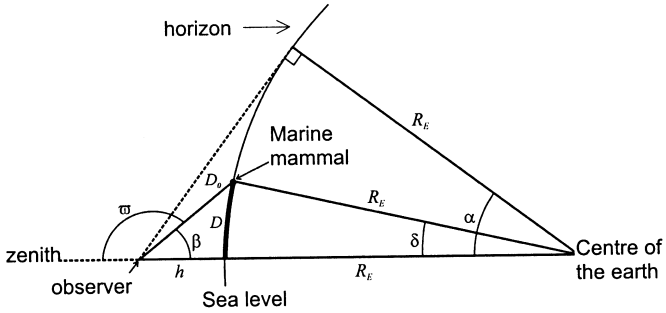


Fig. 1. Schematic showing distances and angles for distance measurements in which the angular drop from the zenith to an object at the sea surface (ϖ or theodolite vertical angle reading) is measured. α is the central arc angle from the horizon to the observation site; β is the angle from the object being fixed to the observation site; and δ is the central arc angle from the object being fixed to the observation site. R_E is the radius of the Earth. The parameter h is the observation site height above sea level. D_0 is the line-of-sight distance to the object and D is the distance to the object along the surface of the surface of the Earth (modified from Lerczak and Hobbs, 1998).

$$\beta = \pi - \varpi$$

$$D_0 = \left(R_E + h\right)\cdot\cos\left(\beta\right) - \sqrt{\left(R_E + h\right)^2\cdot\cos\left(\beta\right)^2 - \left(2hR_E + h^2\right)}$$

$$\delta = \arcsin\left(\sin\left(\beta\right)\frac{D_0}{R_E}\right)$$

$$D = \delta\cdot R_E$$

where:

α = central arc angle from horizon to station;
β = angle from object being fixed to station;
δ = central arc from object being fixed to station;
ϖ = vertical angle estimated with the theodolite (from zenith to object being fixed);
h = theodolite eyepiece height above sea level;
R_E = radius of the Earth (6.371×10^6 m);
D_0 = line-of-sight distance to object being fixed;
D = distance to object being fixed along the surface of the earth/ocean.

Once the distance to the object along the surface of the ocean (D) is known, the great circle equation is used to determine geographic position of the fixed object.

$$\tau = \eta - \rho$$

$$Lat_F = \sin^{-1}\left(\begin{array}{l}\cos\left(\tau\right)\cdot\sin\left(D/60/1852\right)\cdot\cos\left(Lat_S\right)+\\ \left[\sin\left(Lat_S\right)\cdot\cos\left(D/60/1852\right)\right]\end{array}\right)$$

$$Lon_F = \cos^{-1}\left(\frac{\cos\left(D/60/1852\right)-\left[\sin\left(Lat_S\right)\cdot\sin\left(Lat_F\right)\right]}{\cos\left(Lat_S\right)\cdot\cos\left(Lat_F\right)}\right)+Lon_S$$

where:

D = distance in meters between the two points along the surface of the Earth;
τ = bearing from station to object;
η = azimuth or horizontal angle estimated with the theodolite;
ρ = reference azimuth (bearing from station to reference point);
Lat_S = latitude of the station;
Lon_S = longitude of the station;
Lat_F = latitude of the fixed object;
Lon_F = longitude of the fixed object.

The great circumference equation is also used to determine distance between two geographic points along the surface of the Earth when the geographic coordinates (latitude and longitude) of both points are known.

$$D = \left(60\cdot\cos^{-1}\left[\begin{array}{l}\left(\sin\left(Lat_1\right)\cdot\sin\left(Lat_2\right)\right)+\\ \left(\cos\left(Lat_1\right)\cdot\cos\left(Lat_2\right)\right)\cdot\\ \cos\left(Lon_2 - Lon_1\right)\end{array}\right]\right)\cdot1852$$

$$\varphi = \cos^{-1}\left[\frac{\sin\left(Lat_2\right)-\left[\sin\left(Lat_1\right)\cdot\cos\left(D/60\right)\right]}{\sin\left(D/60\right)\cdot\cos\left(Lat_1\right)}\right]$$

where:

D = distance in meters between the two points along the surface of the Earth;
φ = bearing from point 1 to point 2;
Lat_1 = latitude of point 1;
Lon_1 = longitude of point 1;
Lat_2 = latitude of point 2;
Lon_2 = longitude of point 2.

The algorithms were tested using the examples provided by Lerczak and Hobbs (1998) and during subsequent field tests. An example of the real-time distance output is given in Table 1.

Group dispersion data are collected with a series of four 'fixes' (i.e. right-left, front-back). The area of a group is estimated in the shape of a quadrilateral. Although the area occupied by cetacean groups is often other than a quadrilateral, estimating it with only four fixes saves valuable time in the field. This is an important aspect as theodolite fixes used to estimate group dispersion should be taken in the shortest possible time, especially for groups that move quickly.

Data management

Data collected by *Pythagoras* are stored in a *Microsoft Access* database. Each data type (i.e. fix, environmental, focal behaviour, etc.) contains separate tables with relevant data stored in columns within each table. These data are

Table 1

Pythagoras' output of objects fixed with a theodolite at a station (29°19'03.8"N, 94°45'04.8"W) with a station height of 43.189m, eyepiece height of 0.77m and reference azimuth of 79.88°.

Date	Time	Fix type	Group	Behaviour	Declination	Horizontal	Latitude	Longitude	Distance (m)	Bearing
3 June 2000	13:11:28	Vessel	3	Moving	91°40'50"	40°22'20"	29.31090N	94.73791W	1504.33	120.25
3 June 2000	13:13:00	Vessel	3	Moving	91°46'50"	51°40'20"	29.30925N	94.74036W	1419.17	131.55
3 June 2000	13:14:07	Vessel	3	Moving	91°50'20"	54°16'40"	29.30959N	94.74065W	1373.81	131.16
3 June 2000	13:15:10	Vessel	3	Moving	91°45'30"	61°24'50"	29.30763N	94.74205W	1437.25	141.29
3 June 2000	13:16:08	Vessel	3	Moving	91°45'40"	63°23'10"	29.30737N	94.74246W	1434.96	143.27
3 June 2000	13:17:16	Vessel	3	Moving	91°43'30"	67°23'40"	29.30663N	94.74315W	1465.25	147.27
3 June 2000	13:18:46	Vessel	3	Moving	91°46'00"	70°34'50"	29.30652N	94.74404W	1430.41	150.46
3 June 2000	13:20:06	Vessel	3	Moving	92°18'00"	74°11'50"	29.30885N	94.74637W	1096.83	154.08
3 June 2000	13:20:51	Vessel	3	Moving	93°20'40"	73°15'10"	29.31168N	94.74781W	753.00	153.13
3 June 2000	13:21:42	Vessel	3	Moving	95°52'40"	71°30'20"	29.31435N	94.74921W	427.14	151.39
3 June 2000	13:22:56	Vessel	3	Moving	98°07'40"	59°57'00"	29.31561N	94.74927W	307.85	139.83
3 June 2000	13:23:13	Vessel	3	Moving	98°17'40"	63°07'10"	29.31556N	94.74945W	301.58	143.00

visually displayed to the user with *Excel*-like spreadsheets. Search and sort functions further structure and manage the database and data partitioning functions separate records by day, fix type and group. Fix data can be recalculated *post hoc* with updated values in relation to distance and geographic location. Data can be exported to various GIS and database management type files. For further spatial analysis and graphical display, latitude and longitude trackline information can be exported to GIS related files such as *ArcInfo*, *MapInfo*, *Surfer* and *MATLAB*. For additional statistical analysis, data can be saved as *Microsoft Excel*, *Microsoft Access*, or delimited (ASCII) text files.

Analysis

An analysis module was developed to estimate trackline distance and bearing both within and between tracks. For a single trackline, the program calculates leg speed, linearity and reorientation rate per trackline. Leg speed is estimated by calculating the distance travelled between two sequential points within a trackline divided by the time interval between the two points. Linearity is the deviation of a trackline from that of a straight line and is calculated by dividing the net geographic distance between the first and last fix of a trackline by the cumulative distances along the track. Linearity values range between 0 and 1, where a linearity score close to one represents a straight trackline and a value close to zero represents a track with little or no observed directional movement (Batschelet, 1981). Reorientation rates represent a magnitude of bearing changes along a trackline. This rate is calculated as the summation of absolute values of all bearing changes along a trackline divided by the entire duration of the trackline in minutes (Smultea and Würsig, 1995).

Although leg speed, linearity and reorientation rates provide valuable information regarding movement patterns, each of these parameters are limited to information collected within a trackline. Often, theodolite studies that evaluate potential impacts of human-related activities on marine mammals are interested in how animals orient themselves in respect to boats or other moving objects. *Pythagoras* has distance and course estimation modules that determine relationships between tracklines or between a trackline and a fixed point. Comparing tracklines can be difficult due to temporal dependency and the logistic difficulty of fixing two objects at the same time. Algorithms were developed that calculate the distances between two or more tracklines within a defined critical time of actual fixed data or interpolated points. To temporally and spatially interpolate trackline data, the calculated speed and bearing between trackline points were used to estimate the geographic position of the object at specified time intervals within a trackline (Fig. 2). This assumes that objects travel at the same speed and direction between the fixed data points. Since this assumption is not always true for cetaceans, users must specify a critical time-interval (CTI) for interpolation between two consecutive fixes. If the time interval between two consecutive fixes is longer than the critical time, no interpolation will be performed. Defining a critical time interval should take into account the object/species being tracked, its behaviour and other factors that may affect movement. Cipriano (1992) used a 130 sec CTI and Barr (1997) a 240 sec CTI for dusky dolphin studies and Bejder (1997) used 60 sec CTI for research on Hector's dolphins (*Cephalorhynchus hectori*).

Relative orientation is estimated with the scheme devised by Bejder (1999) to interpret directional movements of one object in relation to another. This allows for a quantifiable description of approaches and avoidances between objects and can be used to access potential impacts of anthropogenic activity on marine mammals.

Pythagoras has a number of functions that allow for analysis of behavioural data. For each behavioural event recorded, the defined behaviour, fix type and associated group information are linked by date and time. The program analyses the frequency of behaviours, behavioural time intervals (i.e. blow interval), time between two behavioural events or duration of one event (i.e. the time of 'first surface' to the time of 'dive' = surface time) and rate of a specified behaviour (i.e. blows per minute).

Comparison of theodolite computer systems

To evaluate the benefit that various theodolite programs contribute to cetacean research and to give a basis for future developments, several different systems available for cetacean theodolite tracking were compared. The programs evaluated here are Wolitzky's program (Würsig, 1978), *T-Trak* (Cipriano, 1990), *Aardvark*, *Cyclops* (Kniest *et al.*, 2000) and *Pythagoras* (Gailey and Ortega-Ortiz, 2000). The functionality of each program is presented in four basic categories: (1) general system design; (2) data collection; (3) data management; and (4) analysis (Table 2).

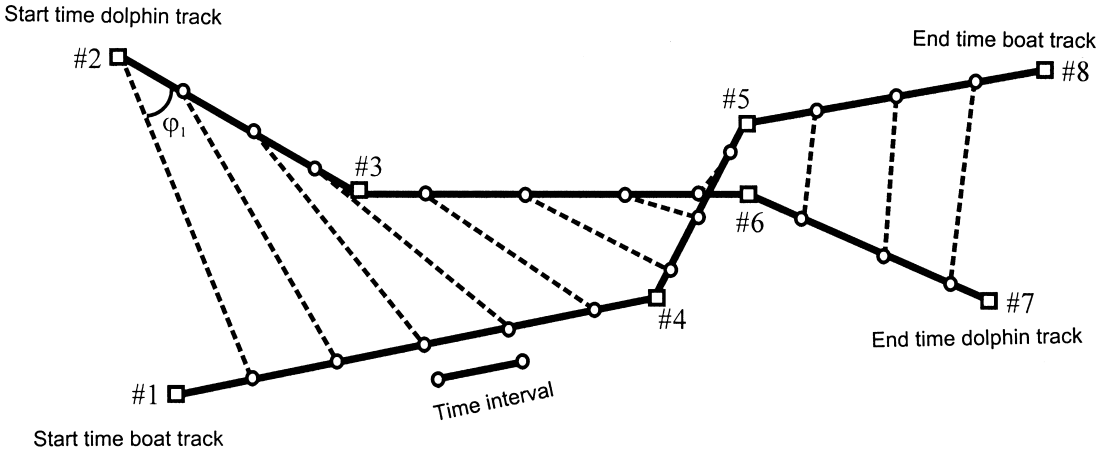


Fig. 2. Positions used to estimate distance between a dolphin trackline and a boat trackline. Location is estimated by interpolating position at specified time intervals. Numbers indicate the sequence of actual fixes. The angle φ indicates the relative orientation of reference trackline (dolphin) to trackline selected for comparison (boat) at time interval one.

Table 2

Comparison of five different theodolite programs. An × indicates the programs ability to perform a function.

Function/feature	Wolitzky's Program	T-Trak	Aardvark	Cyclops	Pythagoras
General					
Operating system	HP Desk Calculator	MS-Dos	Mac	Windows	Windows
Program language	Basic	C		Fortran	Visual Basic
User manual		×		×	×
Supported theodolites			Topcon, Sokkia	Leica, Sokkia, Nikon, Topcon	Topcon, Sokkia
Coordinate system	Cartesian	Cartesian	Cartesian or Lat./Long.	UTM or Lat./Long.	Lat./Long.
Program's database			Text	Binary	Microsoft Access
Data collection					
Data collection program			×	×	×
Environmental			×	×	×
Non-fix data			×		×
Focal behaviour			×		×
Observer data			×		×
Group dispersion				×	×
Real-time display				×	×
GIS maps				Raster	Vector
Real-time calculations				×	×
Data management					
Sorting		×			×
Searching					×
Data partitioning					×
Recalculation		×		×	×
Post-hoc visualisations			×	×	×
Import formats		Text		Text	Excel, Access, Text
Export formats		Text	Text		Excel, Access, Text, ArcInfo, MATLAB, MapInfo, Surfer
Analysis					
Distance calculation	×	×	×	×	×
Leg speed	×	×	×	×	×
Linearity			×		×
Reorientation rate			×		×
Interpolated tracklines			×		×
Relative orientation			×		×
Behaviour					×

The common component of all the programs is the ability to use the required station parameters to convert theodolite vertical and horizontal angles into Cartesian (*x*-*y*) coordinates and/or geographic positions. Although such information is the primary goal of using a theodolite as a research tool, it demonstrates the main benefits of using computer programs to perform calculations based on multiple variables that are constantly changing. Of the five programs, three (*Aardvark*, *Cyclops* and *Pythagoras*) provide a means of collecting data from a computer-connected theodolite in the field. This allows for both accurate and rapid recording of the angles measured by the instrument and increases the number of fixes per trackline. By decreasing the time needed to record each fix, the resolution of the data is increased, not only for the current cetacean of interest, but also for the multitude of other

objects (e.g. boats, swimmers, oil platforms, sources of underwater noise, etc.) that might be relevant to a particular study.

Cyclops and *Pythagoras*, the two most recently developed programs, provide the ability of real-time calculations and trackline display. This is mainly due to the improvement of computer systems and the increased availability of data from GIS databases. Displaying the fixed object position in real-time allows for rapid detection and correction to errors that may occur during a session.

All programs, with the exception of Wolitzky's program, manage data in some form or another. However, *Pythagoras* alone dedicates a separate module designed specifically for database management. *Pythagoras* also has the ability to interact with *Microsoft Access* and *Excel*, two programs commonly used for database management.

Aardvark and *Pythagoras* provide the most detailed analysis routines of the five programs evaluated here. Both programs can calculate leg speed, linearity and reorientation rates within a trackline and distance and relative orientation between two or more tracklines. Behavioural data analysis is only available in *Pythagoras*.

DISCUSSION

Although theodolites are highly accurate in determining geographic locations of fixed objects, errors can always present a problem in calculations. Würsig *et al.* (1991) described asymmetrical errors caused by inaccurate estimates of observation platform height. The estimated error decreases with increasing elevation of the observation height, and increases with increasing distance of the object from the theodolite. Swell height produces errors due to the change in height of an object being fixed at the crest or trough of a swell. As object distances increase with respect to the station, refraction of light affects line of sight estimation.

The analysis of movement patterns presented here can be improved by incorporating algorithms to evaluate tracklines in terms of correlated random walks (CRW) as suggested by Turchin (1998). The use of CRW models can provide estimates of diffusion rates of individuals and the system can evaluate the appropriateness of CRW models with net squared displacement plots (Turchin, 1998). Other analytical considerations for future system developments should include distance-to-shore, depth profile, and habitat use analysis, which can increase the efficiency in analysing theodolite data.

As computers increase in computational power, increasingly sophisticate software programs are being developed for cetacean research, and often target a specific species, type of data to be collected and analytical approach. Individual photographic identification, for instance, has been aided by the development of computer-assisted recognition systems (Hiby and Lovell, 1990; Mizroch *et al.*, 1990; Whitehead, 1990; Gailey, 2001). Similarly, population estimates and analysis of social patterns have benefited greatly from other programs (Menkens and Anderson, 1988; Whitehead, 1999).

Relatively few such programs exist for theodolite-based research. The system described here provides researchers with a tool to efficiently collect, manage and analyse theodolite-obtained cetacean movement and behaviour data in the field. It also provides functions that can benefit cetacean studies and allow easier integration of data to and from other tools (e.g. database management, GIS, GPS). With the increasing concern of human-related impact on marine mammal populations, quick, efficient and accurate processing of theodolite-based data may allow for rapid assessment of potential impacts, and timely responses to possible management issues.

ACKNOWLEDGEMENTS

We are grateful to Bernd Würsig for his support towards the development of *Pythagoras*. We appreciate Adam Frankel's assistance in supplying previous source codes of theodolite programs. We also thank Lars Bejder, Lisa Schwarz and Sylvie Rimella for their helpful suggestions. Alice Mackay spent endless hours trying to find potential bugs in the program. The manuscript was improved by comments from Bernd Würsig, Dave Weller, Leszek Karczmarski, Suzanne Yin and two anonymous reviewers. Financial support for JGO was provided by a fellowship from CONACyT, Mexico. This represents Contribution No. 78 of the Marine Mammal Research Program, Texas A&M University at Galveston.

REFERENCES

Barr, K. 1997. The impacts of marine tourism on the behaviour and movement patterns of dusky dolphins (*Lagenorhynchus obscurus*) at Kaikoura, New Zealand. M.S. Thesis, University of Otago, Dunedin, New Zealand. 97pp.

Batschelet, E. 1981. *Circular Statistics in Biology*. Academic Press, London. 371pp.

Bejder, L. 1997. Behaviour and ecology of Hector's dolphins (*Cephalorhynchus hectori*) in Porpoise Bay, New Zealand and the impacts of tourism. M.S. Thesis, University of Otago, Dunedin, New Zealand. 101pp.

Bejder, L., Dawson, S. and Harraway, J.A. 1999. Responses by Hector's dolphins to boats and swimmers in Porpoise Bay, New Zealand. *Mar. Mammal Sci.* 15(3):738-50.

Brown, N.C. 2000. The dusky dolphin, *Lagenorhynchus obscurus*, off Kaikoura, New Zealand: the long-term comparison of behaviour and habitat use. M.S. Thesis, University of Auckland, New Zealand. 163pp.

Cipriano, F. 1992. Behavior and occurrence patterns, feeding ecology, and life history of dusky dolphins (*Lagenorhynchus obliquidens*) off Kaikoura, New Zealand. Doctoral Thesis, University of Arizona. 216pp.

Cipriano, R. 1990. *User's Manual T-TRAK Version 1.0. Theodolite-Tracking Data Analysis (for IBM-compatible Computers)*. University of Arizona, Arizona, USA. 38pp.

Gailey, G. 2001. Computer systems for photo-identification and theodolite tracking of cetaceans. M.S. Thesis, Texas A & M University, College Station. 125pp.

Gailey, G. and Ortega-Ortiz, J.G. 2000. *User's Manual 'Pythagoras' Version 1.0: Theodolite cetacean tracking*. Texas A & M University, Galveston, USA. [Available at: *http://www.tamug.tamu.edu/mmrp/pythagoras*].

Hiby, A.R. and Lovell, P. 1990. Computer aided matching of natural markings: a prototype system for grey seals. *Rep. int. Whal. Commn* (special issue) 12:57-61.

Kniest, E., Baker, G.S. and Suchanow, W. 2000. *User's Manual 'CYCLOPS' Version 3.1*. University of Newcastle, Australia. [Available at: *http://mimas.newcastle.edu.au/~cyclops*].

Lerczak, J.A. and Hobbs, R.C. 1998. Calculating sighting distances from angular readings during shipboard, aerial, and shore-based marine mammal surveys. *Mar. Mammal Sci.* 14(3):590-9. [See Errata. 1998. *Mar. Mammal Sci.* 14(4):903].

Martin, B. and Bateson, P. 1993. *Measuring Behaviour*. 2nd Edn. Cambridge University Press, New York. 222pp.

Menkens, G.E., Jr. and Anderson, S.H. 1988. Estimation of small-mammal population size. *Ecology* 69:1952-9.

Mizroch, S.A., Beard, J. and Lynde, M. 1990. Computer assisted photo-identification of humpback whales. *Rep. int. Whal. Commn* (special issue) 12:63-70.

Smultea, M.A. and Würsig, B. 1995. Behavioural reactions of bottlenose dolphins to the *Mega Borg* oil spill, Gulf of Mexico 1990. *Aquat. Mamm.* 21(3):171-81.

Turchin, P. 1998. *Quantitative Analysis of Movement*. Sinauer Associates, Sunderland, Massachusetts. 396pp.

Ward, B.G. 1999. Movement patterns and feeding ecology of the Pacific coast bottlenose dolphins (*Tursiops truncatus*). M.Sc Thesis, San Diego State University, San Diego, California. 98pp.

Whitehead, H. 1990. Computer assisted individual identification of sperm whale flukes. *Rep. int. Whal. Commn* (special issue) 12:71-7.

Whitehead, H. 1999. *Users's Manual 'SOCPROG' version 1.3*. Dalhousie University, Halifax, Nova Scotia, Canada. [Available at: *http://is.dal.ca/~hwhitehe/social.htm*].

Würsig, B. 1978. On the behavior and ecology of bottlenose and dusky porpoises in the south Atlantic. Ph.D. Thesis, State University of New York at Stony Brook. 315pp.

Würsig, B., Cipriano, F. and Würsig, M. 1991. Dolphin movement patterns: information from radio and theodolite tracking studies. pp. 79-111. *In:* K. Pryor and K.S. Norris (eds.) *Dolphin Societies, Discoveries and Puzzles*. University of California Press, Berkeley, California, USA. 397pp.

Yin, S.E. 1999. Movement patterns, behaviors and whistle sounds of dolphin groups off Kaikoura, New Zealand. M.S. Thesis, Texas A & M University, College Station. 107pp.