

Quantifying Uncertainty and Incorporating Environmental Stochasticity in Stock
Assessments of Marine Mammals

John Robert Brandon

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

University of Washington

2009

Program Authorized to Offer Degree:

School of Aquatic and Fishery Sciences

University of Washington
Graduate School

This is to certify that I have examined this copy of a doctoral dissertation by

John Robert Brandon

and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by the final
examining committee have been made

Chair of the Supervisory Committee:

André E. Punt

Reading Committee:

André E. Punt

Timothy E. Essington

Paul R. Wade

Judith E. Zeh

Date: _____

In presenting this thesis in partial fulfillment of the requirements for the doctoral degree at the University of Washington, I agree that the Library shall make its copies freely available for inspection. I further agree that extensive copying of the dissertation is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for copying or reproduction of this dissertation may be referred to ProQuest Information and Learning, 300 North Zeeb Road, Ann Arbor, MI 48106-1346, 1-800-521-0600, to whom the author has granted "the right to reproduce and sell (a) copies of the manuscript in microform and/or (b) printed copies of the manuscript made from microform."

Signature_____

Date_____

University of Washington

Abstract

Quantifying Uncertainty and Incorporating Environmental Stochasticity in Stock Assessments of Marine Mammals

John Robert Brandon

Chair of the Supervisory Committee:
Professor André E. Punt
School of Aquatic and Fishery Sciences

Incorporating scientific uncertainty and accounting for potential effects of environmental variability in stock assessments are vital aspects of providing sound management advice. This dissertation focuses on these issues in the context of estimating sustainable quotas for aboriginal subsistence hunting of marine mammals.

Bayesian model averaging was used to take into account both parameter and model uncertainty in the assessment of the Bering-Chukchi-Beaufort Seas (BCB) stock of bowhead whales. The lower 5th percentile of the Bayesian model-averaged posterior for an aboriginal whaling catch quantity (Q_1^{+}) was estimated to be 155 whales in 2002. This estimate provides confirmatory evidence that current catch quotas for this stock are sustainable.

In order to assess the robustness of Bayesian assessments to alternative methods for constructing a joint prior distribution which respects biological realism, sensitivity analyses were performed in the context of a risk assessment. The probability of meeting a management objective for aboriginal subsistence hunting was found to be robust for the data-rich BCB bowhead example. However, the data-poor East Greenland walrus example was shown to be sensitive to this issue.

A framework was developed and applied to the assessment of the eastern North Pacific stock of gray whales which incorporated environmental stochasticity by estimating the relationship between residuals in birth and survival rates to an environmental time series. The scenarios which accounted for the effect of the mortality event in 1999 and 2000 led to less optimistic estimates of population status during recent decades with concomitant recovery generally attributable to higher survival rates as opposed to higher birth rates. This framework was then used as an operating model conditioned on forecasts of sea-ice, which has been hypothesized to regulate recent calf production in the population. Future population trajectories were then used to test the performance of the current catch control rule for this stock. Performance was measured by the ability to satisfy different levels of future aboriginal subsistence need, and was shown to be satisfactory under the range of trials examined here.

The methods developed in this dissertation are generally applicable for providing management advice for exploited populations of marine mammals and other renewable natural resources.

TABLE OF CONTENTS

	Page
List of Figures	ii
List of Tables.....	iii
Introduction	6
Chapter 1: Assessment of the Bering-Chuckchi-Beaufort Seas stock of bowhead whales using Bayesian model averaging.....	12
1.1 Introduction.....	13
1.2 Methods.....	17
1.3 Results.....	27
1.4 Discussion	33
Chapter 2: Constructing a coherent joint prior while respecting biological realism: sensitivity of marine mammal stock assessments to recent methods.....	53
2.1 Introduction.....	54
2.2 Methods.....	58
2.3 Results.....	68
2.4 Discussion	72
Chapter 3: Assessment of the eastern stock of North Pacific gray whales: incorporating calf production, sea-ice and strandings data.....	95
3.1 Introduction.....	96
3.2 Methods.....	98
3.2 Results.....	113
3.3 Discussion	117
Chapter 4: Testing the Gray Whale <i>SLA</i> : allowing environmental variability to influence population dynamics	136
4.1 Introduction.....	136
4.2 Methods.....	138
4.3 Results.....	145
4.4 Discussion	147
Conclusions.....	158
References	165
Appendix A: Derivation of Eqns. 3.2 and 3.3.....	181

LIST OF FIGURES

Figure Number	Page
1.1 Model fits to abundance estimates for BCB bowheads	47
1.2 Posterior distributions for carrying capacity and 2002 abundance	48
1.3 Posterior distributions for depletion and intrinsic growth rate.....	49
1.4 Posterior distributions for replacement yield and Q_1^{1+}	50
1.5 Posterior distributions for juvenile and adult survival rates.....	51
1.6 Posterior distributions for age at maturity and maximum fecundity.....	52
2.1 Contour plots showing solutions for S_{juv} for EG walrus	85
2.2 Model fits and estimated abundance for BCB bowheads	86
2.3 Priors and posteriors for K and current depletion for BCB bowheads.....	87
2.4 Priors and posteriors for a_m and f_{max} for BCB bowheads	88
2.5 Priors and posteriors for S_a and S_{juv} for BCB bowheads	89
2.6 Priors and posteriors for λ_{max} for BCB bowheads.....	90
2.7 Trajectories of 1+ abundance for BCB bowheads	91
2.8 The probability of meeting a management objective.....	92
2.9 Trajectories of total population size for EG walrus	93
2.10 Bivariate scatterplots of parameter values	94
3.1 Life cycle graph for ENP gray whale model.....	130
3.2 Sea-ice (HadSST) index and catches for ENP gray whales.....	130
3.3 MCMC diagnostics for ENP gray whales.....	131
3.4 Model fits for ENP gray whales.....	132
3.5 Sensitivity analyses for ENP gray whales.....	133
3.6 Posteriors for depletion and K for ENP gray whales	134
3.7 Posterior for birth and survival rates for ENP gray whales	135
4.1 Life cycle graph used to test Gray Whale <i>SLA</i>	154
4.2 Sea-ice forecast and future catastrophes used to test Gray Whale <i>SLA</i>	155
4.3 Trajectories and future catches from testing Gray Whale <i>SLA</i>	156
4.4 Time trajectories of future process error residuals.....	157

LIST OF TABLES

Table Number	Page
1.1 BCB bowhead abundance estimates	37
1.2 Proportion of observed calves and mature animals, 1985-1994	38
1.3 Catches of BCB bowhead whales, 1848–2002	39
1.4 Prior distributions for BCB bowheads	40
1.5 Posterior distributions for management related quantities	42
1.6 Posterior distributions for life history parameters	44
1.7 Bayes factors for comparisons of paired models	46
1.8 Posterior model probabilities used in Bayesian model averaging	46
2.1 BCB bowhead abundance estimates	80
2.2 Proportion of observed bowhead calves and mature animals, 1985-1994	81
2.3 Prior distributions for BCB bowheads	82
2.4 Prior distributions for EG walrus	83
2.5 Posteriors distributions for BCB bowheads	84
2.6 Posteriors distributions for EG walrus	84
3.1 Scenarios considered for ENP gray whales	123
3.2 Prior distributions for ENP gray whales	125
3.3 Results for the Full model for ENP gray whales	126
3.4 Results when ε_{add-1} & ε_{add-2} are estimated individually for ENP gray whales	128
4.1 Scenarios considered for testing the Gray Whale <i>SLA</i>	149
4.2 Performance statistics results for testing the Gray Whale <i>SLA</i>	150

ACKNOWLEDGEMENTS

I thank André Punt for being my advisor and major professor. Through this process I have come to trust him completely and am honored to have had the opportunity to be one of his students. His dedication and passion are unsurpassed. There may be other advisors which are as good, but I am certain there can be none better.

My committee members have provided outstanding feedback and a critical eye on my research. Paul Wade was the reason I came to Seattle, and he sat down with me during the first phase of this research and patiently helped me cut my teeth on some very challenging problems. Just as importantly however, he has opened many doors in my life outside of this research. Every time I have gone to Paul for advice, he has been there and I have walked away better for it. Judy Zeh possesses the rare combination of statistical expertise and pragmatism that make her the consummate problem solver. I have been fortunate to have seen her in action in the Scientific Committee of the IWC; when Judy speaks everyone listens because the solution to the problem is not far behind. Tim Essington has provided a valuable breadth of perspective to my committee. More than anyone, he has challenged me to see the big picture. This is something I hope to continue to struggle with (in a good way) for the rest of my career. Eliot Brenowitz served as my GSR and generously took time away from his own research to join us in meetings whenever the occasion arose.

The highlight of this research has been working with and learning alongside other graduate students, especially my lab-mates: Teresa A'mar, Jason Cope, Gavin Fay, Melissa Haltuch, Doug Kinzey, Tommy Garrison, Rod Towell, Chantel Wetzel and Motoki Wu.

Many people have provided reviews and ideas during this process. I would like to thank members of the Scientific Committee of the IWC for their feedback, especially: Jeff Breiwick, Robert Brownell Jr., Doug Butterworth, Justin Cooke, Geof Givens, Bill Koski, Sue Moore and Lars Witting. I would also like to thank Dave Rugh and several anonymous reviewers for their reviews on the published chapters. Likewise, this work would not be possible were it not for the many colleagues in the field who have collected and analyzed the data used in the analyses. I thank Muyin Wang for providing the global climate model forecasts of sea-ice. I would also like to extend my thanks to Wayne Perryman for his help with my research on ENP gray whales.

Financial support was provided by Washington Sea Grant, University of Washington, pursuant to the National Oceanic and Atmospheric Administration Graduate Fellowship Program in Population Dynamics and Marine Resource Economics, and also through the Minerals Management Service, the IWC and the School of Aquatic and Fishery Sciences.

Finally, I would like to thank Tim Gerrodette for getting my foot in the door and my wonderful girlfriend Janna Wemmer for her love and support along the way.

DEDICATION

To my father and mother, John and Dorothy Brandon.

Thank you.

INTRODUCTION

Incorporating scientific uncertainty and accounting for potential effects of environmental variability in stock assessments are vital aspects of providing sound management advice (Punt and Hilborn, 1997; Maunder and Watters, 2003). This dissertation focuses on these issues in the context of estimating sustainable catch quotas for aboriginal subsistence hunting of marine mammals. Case studies utilize data for three stocks subject to hunting: the Bering-Chukchi-Beaufort seas (BCB) stock of bowhead whales (*Balaena mysticetus*), walrus off East Greenland (EG) (*Odobenus rosmarus rosmarus*), and the eastern North Pacific (ENP) stock of gray whales (*Eschrichtius robustus*). Specifically, this work advances existing Bayesian stock assessment methods by examining and applying methods which: 1) account for model uncertainty through Bayesian model averaging; 2) evaluate the sensitivity of assessment results to alternative approaches for constructing coherent joint prior distributions which respect biological realism; 3) synthesize information relating environmental variables and population processes, and 4) use simulations to test the robustness of an existing management procedure given climate forecasts and plausible scenarios for how population dynamics might be affected by such. The overarching theme is incorporation of uncertainty at different levels (e.g. parameter, observation, process, and model uncertainty), to better represent and improve the current understanding of population dynamics and resulting estimates of sustainable catch quotas, hence providing the best available management advice.

A great deal of progress has been made during the last decade to better quantify and integrate various sources of uncertainty which are inherent in calculating sustainable catch quotas. The Scientific Committee of the International Whaling Commission (IWC) has played an instrumental role in the development of Bayesian stock assessment methods to meet this goal. The Bayesian methods that were pioneered by the IWC were initially in response to the challenge of providing management advice on catch quotas for the aboriginal hunt of BCB bowhead

whales, given uncertainty surrounding biological parameters and stock abundance (Raftery *et al.*, 1995). Subsequent extensions of that approach have formed a widely accepted paradigm for integrating a variety of sources of information and concomitant uncertainty into fisheries stock assessments (Punt and Hilborn, 1997). However, one aspect of uncertainty that is still largely underestimated is model uncertainty (Clyde and George, 2004). While model selection has become an increasingly more popular tool (Burnham and Anderson, 2002), it is not always advisable to base inference on a single model or hypothesis. This is especially true when ambiguity exists with respect to which model best fits available observations. Indeed, this is exactly the case for the BCB bowheads; different population dynamics models are able to fit the data nearly equally well, yet the resulting estimates for quantities of management interest differ substantially among models. Model averaging provides an attractive technique for dealing with such a conundrum. Therefore, the first chapter of this dissertation expands upon earlier IWC assessments of the BCB stock of bowhead whales by explicitly accounting for this additional dimension of uncertainty, through Bayesian model averaging.

Bayesian analysis allow for various sources of prior knowledge (or ignorance), to be combined into a joint prior distribution, upon which inference may proceed for estimated parameters and derived quantities. However, some aspects of prior information may be inconsistent with each other. Such inconsistencies are likely to be more common for models that contain many estimated parameters that are functionally related. Bayesian stock assessments of marine mammals using deterministic age-structured population dynamics models are an illustrative example of this problem. In this context, life history parameters and the parameter for the population growth rate are functionally related, which raises two related issues: 1) placing explicit priors on each life history parameter as well as on the population growth rate parameter results in an incoherent joint prior distribution (i.e. two different priors on the estimated parameters), and; 2) certain combinations

of values drawn from the priors may result in biologically implausible solutions for the remaining parameter (which can be solved for analytically, to form a coherent joint prior, given the functional relationship mentioned above). Different approaches have been applied in recent assessments of marine mammals to deal with this problem and the issue of how best to satisfy constraints on parameter values given the realm of biological plausibility. However, the sensitivity of assessment results to this issue has not been explored in any detail to date. Thus, the second chapter of this dissertation investigates the sensitivity of assessment and risk analysis results, using data-rich (BCB bowhead) and data-poor (EG walrus) case studies for comparison, to better understand the potential consequences of these alternative approaches on the estimates of sustainable catch quotas.

It is becoming increasingly more certain that climate is changing at unprecedented rates (IPCC, 2007). Stock assessments of marine mammal populations conducted by the U.S. National Marine Fisheries Science Service (NMFS) and the IWC are generally based on age- and sex-structured population dynamics models, with density dependence acting on fecundity (Breiwick *et al.*, 1984; Punt, 1999a; Wade, 2002a). An underlying assumption of these models is that fecundity is related deterministically to the size of some component of the population, and is independent of changes in the physical environment. However, the number of studies providing evidence that environmental variability (e.g. timing and extent of sea-ice break-up) is correlated with female foraging and reproductive success are growing. Examples include: polar bears (e.g., Stirling *et al.* 1999), northern and Antarctic fur seals (Goebel, 2002; Forcada *et al.*, 2005), northeast Atlantic fin whales (Lockyer, 1986), north and south Atlantic right whales (Green *et al.*, 2003; Leaper *et al.*, 2006), sperm whales (Whitehead, 1997) and ENP gray whales (Perryman *et al.*, 2002). Simmonds and Isaac (2007) review expected effects of climate change on a variety of marine mammal species,

highlighting, "...the need to take projected impacts into account in future conservation and management plans, including species assessments."

Accordingly, management strategies should strive to be robust to a range of plausible consequences of climate change. This issue is of immediate importance to stocks of marine mammals in the arctic and sub-polar oceans, because it is in these areas of the world where the effects of climate change are likely to be most profound (ACIA, 2004). For example, climate forecasts predict a massive decrease in the extent of arctic sea ice (Overland and Wang, 2007). This change is expected to have major impacts on populations of marine mammals which live in or migrate to the arctic, many of which are also important to aboriginal subsistence and culture (Lowry, 2000). Consequently, understanding the biological influences of environmental variability, especially with regards to marine mammals in the arctic and sub-polar oceans, is an area of special concern to both conservation and management. Hence, the third chapter of this dissertation advances existing stock assessment methods for marine mammals by investigating a framework to synthesize information relating environmental variables and population processes. The goal of this work is to provide improved estimation and evaluation of management quantities and strategies given natural variability in ecosystem conditions, and forecasts of climate change. The ENP gray whale is used as a case study because this stock is currently the target of aboriginal subsistence hunting and there exists suitable data to complete the objectives of this research.

While stock assessments are an important aspect of monitoring recoveries and understanding population dynamics, management procedures are increasingly being adopted as a way for setting catch quotas to achieve sustainability and conservation goals (Butterworth, 2007; Punt and Donovan, 2007). A fundamental component of the selection of a management procedure is the use of computer simulations to examine the performance of a proposed catch control rule (for aboriginal subsistence whaling the IWC catch control rule is referred to as a Strike

Limit Algorithm, *SLA*, and determines the number of allowable “strikes” each year). Underlying the simulation testing structure is an operating model, which takes the major sources of identified uncertainty (e.g., population growth rate, stock structure, etc.) into account. The operating model corresponds to a hypothesized underlying ‘true’ state of nature and is used to generate the simulated data that is provided to the management procedure. Given this information, the strike (or catch) limit is computed by the management procedure and the resulting mortality subtracted from the population represented in the operating model. This feedback process is repeated as the underlying population is projected forward through time. At the end of the projection period, performance statistics are compiled based on management criteria. Often, multiple operating models are used during this process to investigate the performance of a proposed management procedure relative to alternative scenarios regarding the ‘true’ state and dynamics of nature. Recommend catch limits for ENP gray whales are based on the ‘Gray Whale *SLA*’, which has undergone extensive simulation testing. However, it would be prudent to revisit the performance of these control rules given the availability of new and detailed forecasts for changes in sea ice, and evidence which suggests that variability in calf production for ENP gray whales may be related to variability in sea ice. Therefore, the fourth chapter of this dissertation involves the development of an operating model (based on the assessment in chapter 3) to evaluate the performance of the Gray Whale *SLA* and hence assess its robustness to environmental variation.

Objectives

The objectives of this dissertation are:

1. perform an assessment of the BCB stock of bowhead whales based on a variety of alternative population dynamics models, which represent competing hypotheses, and use Bayesian model averaging to incorporate model uncertainty in the resulting management advice;

2. perform and summarize the sensitivity of assessment and risk analysis results to alternative approaches for constructing a coherent joint prior while respecting biological realism. Compare the sensitivity of the results between data-rich (BCB bowhead) and data-poor (EG walrus) case studies;
3. develop methods for synthesizing environmental time series with existing stock assessment models for marine mammals, to capture alternative assumptions regarding biological processes (e.g. fecundity is not only related to population size, but also to some function of an environmental variable) and apply the framework to an assessment of the ENP gray whales; and,
4. use the framework developed in chapter 3 as an operating model within a simulation modeling approach, to evaluate the performance of the IWC Gray Whale *SLA* given hypotheses regarding the nature and extent of future environmental stochasticity and forecasts of a relevant climate index.

Chapter 1:

Assessment of the Bering-Chukchi-Beaufort Seas stock of bowhead whales using Bayesian model averaging¹**ABSTRACT**

Bayesian estimation methods are used to fit an age- and sex-structured population model to available data on abundance and stage-proportions (i.e., calves/mature animals in the population) for the Bering-Chukchi-Beaufort Seas stock of bowhead whales (*Balaena mysticetus*). The analyses consider three alternative population modeling approaches: 1) modeling the entire population trajectory from 1848, using the ‘backwards’ method where the trajectory is back-calculated based on assigning a prior distribution to recent abundance; 2) modeling only the recent population trajectory, using the ‘forwards from recent abundance’ method, where the population is projected forwards from a recent year and the abundance in that year is not assumed to be at carrying capacity; and 3) a version of 2) that ignores density-dependence. The ‘backwards’ method leads to more precise estimates of depletion level. In contrast, the ‘forwards from recent abundance’ method provides an alternative way of calculating catch-related quantities without having to assume that the catch record is known exactly from 1848 to the present, or having to assume that carrying capacity has not changed since 1848. Not only are all three models able to fit the abundance data well, but each is also able to remain consistent with available estimates of adult survival and age of sexual maturity. Sensitivity to the stage-proportion data and the prior distributions for the life history parameters indicates that use of the 1985 stage-proportion data has the greatest effect on the results, and that those data are less consistent with data on

¹ This chapter was published in The Journal of Cetacean and Research Management:

Brandon, J. and Wade, P.R. 2006. Assessment of the Bering-Chukchi-Beaufort Seas stock of bowhead whales using Bayesian model averaging. *J. Cetacean Res. Manage.* 8: 225-39.

trends in abundance and age of sexual maturity. The analyses indicate that the population has approximately doubled in size since 1978, and the ‘backwards’ analyses suggest that the population may be approaching carrying capacity, although there is no obvious sign in the data that the population growth rate has slowed. Bayes factors are calculated to compare model fits to the data. However, there is no evidence for selecting one model over another, and furthermore, the models considered in this study result in different posterior distributions for quantities of interest to management. Posterior model probabilities are therefore calculated and used as weights to construct Bayesian model-averaged posterior distributions for outputs shared among models to take this ambiguity into account. This study represents the first attempt to explicitly quantify model uncertainty when conducting a stock assessment of bowhead whales.

1.1 INTRODUCTION

Bowhead whales (*Balaena mysticetus*) have been caught by Alaskan Eskimos in an aboriginal fishery for at least hundreds of years; active whaling started along the arctic coast around 1800-1700 BP (Braham, 1989; Dinesman and Savinetsky, 2003). A commercial fishery began takes from the Bering-Chukchi-Beaufort Seas (B-C-B) stock of bowhead whales in 1848, and the population is thought to have been severely depleted by the substantial commercial catches that occurred during the 1800s (Bockstoce and Botkin, 1983; Breiwick and Braham, 1990; Breiwick *et al.*, 1984). Although the commercial fishery almost completely collapsed early in the 1900’s, aboriginal catches of B-C-B bowheads continue (*e.g.*, Braham, 1995; George *et al.*, 1988; Suydam and George 2004).

Most of the B-C-B bowhead stock migrates seasonally along the north slope of Alaska between wintering areas in the Bering Sea and summer feeding areas in the Beaufort Sea. Surveys have been conducted during the spring migration past Pt. Barrow in eleven years since 1978 to estimate the abundance and trends of this population (George *et al.*, 2004). The ice-based counts have resulted in abundance

estimates substantiated by estimates using mark-recaptures of individually identifiable whales in aerial photographs (da Silva *et al.*, 2000). These abundance estimates have served as the primary basis for assessments of the status of and management advice for the B-C-B bowhead stock by the Scientific Committee of the International Whaling Commission (IWC).

Given a history of catches, it was possible to conduct an initial assessment of the status of the population once the first abundance estimate was made in 1978. This assessment suggested that the population had experienced some recovery since its depletion in the 1800s, but that it was still below carrying capacity (Breiwick *et al.*, 1984). Breiwick and Braham (1990) noted that estimates of carrying capacity and depletion level are sensitive to the current estimated population size. However, it is now possible to make more precise estimates of these quantities (conditioned on the accuracy of the catch data and population dynamics model) because the number and precision of the abundance estimates has increased substantially.

Givens *et al.* (1993) and Raftery *et al.* (1995) introduced Bayesian methods (*e.g.*, Press, 1989) to the assessment of bowhead whales. These methods can make use of multiple sources of data and fully characterise uncertainty. Bayesian assessments of the B-C-B bowhead stock (*e.g.* Givens *et al.*, 1995) using the BALEEN II model (de la Mare, 1989) have consequently been used extensively by the Scientific Committee of the IWC (Punt, 1999a). Although there was general agreement that using multiple sources of data in assessments was desirable, considerable debate ensued in the Scientific Committee of the IWC regarding the details of the statistical methods to be used in bowhead assessments (*e.g.*, Butterworth and Punt, 1995; Punt and Butterworth, 1999, 2000; Poole *et al.*, 1999; Poole and Givens, 2001; Schweder and Ianelli, 2000). One troublesome outcome of the results in the standard ‘forwards from K’ assessment (*i.e.* projecting the population dynamics model forwards from a prior distribution for the population

size in 1848, which was assumed to be carrying capacity) was that the observed rate of increase of the population (ROI) differed substantially from the resulting posterior distribution for the ROI, even though the data on ROI were the main data source used in the analysis (*e.g.* Punt and Butterworth, 1999; Raftery and Poole, 1997). This and other related issues were determined to be due in part to a methodological issue associated with the forwards method and the specification of prior distributions. It was also determined that these issues could essentially be resolved by the use of the ‘backwards’ method, in which no prior is specified for the population size in 1848; instead, a prior distribution is specified for the abundance in a recent year, and the population level in 1848 (assumed to be carrying capacity) is then back-calculated from that recent abundance (Butterworth and Punt, 1995; Punt and Butterworth, 1999; Poole and Raftery, 1998).

An important assumption made in bowhead assessments is that the catch history is known without error. It has been found that the catch record going back to 1842 for eastern North Pacific stock of gray whales cannot be reconciled with the population’s current dynamics (Punt and Butterworth, 2002), without assumptions that the catch record has been substantially under-estimated and/or that carrying capacity has changed since the mid 1800s. While it appears from past assessments that the bowhead catch record can be reconciled with the current dynamics, it is still appropriate to investigate an alternative assessment that does not make use of the historical catch record. This can be achieved using the method developed by Wade (2002a) for gray whales, where a model is projected forwards from the abundance in a recent year, with a separate prior distribution specified for carrying capacity. Such an assessment does not use the historic catch record prior to 1978, and does not need to assume that carrying capacity has remained constant since the mid 1800s. Therefore, such an assessment should be robust to problems with these assumptions.

The 1998 assessment of the B-C-B bowhead stock (IWC, 1999; Punt, 1999b) used ‘backwards’ Bayesian estimation based on the BALEEN II model, and the then-available abundance estimates and the data on the proportion of calves and mature animals in the population during 1985-94 (the ‘stage-proportion’ data). New information available for this stock since the 1998 assessment includes: 1) a mark-recapture estimate of adult survival from aerial photographs (Zeh *et al.*, 2002); 2) an estimated age at sexual maturity and an estimate of maximum age (George *et al.*, 1999); 3) a recalculation of the stage-proportion data (Koski *et al.*, 2006); and 4) an estimate of abundance for 2001 (George *et al.*, 2004).

This chapter examines the sensitivity of the results of the B-C-B bowhead assessment to: 1) modeling the entire population trajectory from 1848 (using the ‘backwards’ method); and 2) modeling only the recent period (where the population is projected forwards from a recent year, and the abundance in that year is not assumed to be at carrying capacity). A third set of analyses is conducted using a Leslie matrix with no density-dependence. The population model used in this assessment is the density-dependent Leslie matrix model (Leslie, 1945, 1948) developed by Breiwick *et al.* (1984) except that reproduction and natural mortality are assumed to occur before removal by catches. When parameterized in equivalent ways, the results of the BALEEN II model and this Leslie matrix model have been found to be nearly identical (Wade, 2002a; Punt and Butterworth, 2002).

This chapter also examines the sensitivity of the model outputs to the sources of data included in the assessment and the specifications for the prior distributions. Scenarios are specified to investigate the sensitivity of the results to data-based informative vs. uniform (less informative) prior distributions, as well as to how the stage-proportion data are constructed. The stage-proportion data are excluded altogether in some analyses, and varied in others based on whether the aerial

photographic survey data from 1985² are included in their calculation or not. Within a scenario, identical prior distributions are assumed for the life-history parameters and matching data-sets are used in the likelihood function for each of the three modeling approaches to enable comparisons to be made among them.

Model uncertainty is accounted for within a Bayesian framework, and Bayes factors (Kass and Raftery, 1995) are calculated to quantify the evidence provided by the data in favour of the different modeling approaches. For the ‘reference scenarios,’ there is no evidence for selecting one model over another, but there are important differences in the estimates of quantities that would be of interest to management. Therefore, we follow the philosophy outlined by a number of authors (e.g., Kass and Raftery, 1995; Buckland *et al.*, 1997; Hoeting *et al.*, 1999; Patterson, 1999; Durban *et al.*, 2005; Raftery *et al.*, 2005) and present quantities of interest as Bayesian model averages; weighting the output of contending models based upon their relative likelihoods, as opposed to selecting between them.

1.2 METHODS

Available data

The population dynamics models are fit to three sources of data: 1) abundance estimates from ice-based surveys at Pt. Barrow, Alaska between 1978 and 2001 (Table 1.1); 2) proportion calves/mature animals in the population from 1985 to 1994 (Table 1.2); and 3) annual catches in individuals from 1848 to 2002 (Table 1.3).

All of these sources of data were used in the 1998 assessment but have been updated since. The first ice-based survey since 1993 was conducted in 2001, and has provided an abundance estimate for that year of 10,545 with a coefficient of variation of 0.128 (updated from George *et al.*, 2004 by Zeh and Punt, 2005). The

² 1985 appears to have been an anomalously late migration year; the aerial survey in 1985 is believed to have ended too early to have sampled the tail-end of the migration, typically when most cow/calf pairs are in the study area (Koski *et al.*, 2006)

catch data have been updated with the post-1998 catches and revisions to the catches for 1994-96. Two additional years of aerial photographic data have been analyzed since the previous assessment, and the stage-proportion data have been recalculated (Angliss *et al.*, 1995; Koski *et al.* 2006).

Additional information available for this stock since the 1998 assessment includes a mark-recapture estimate of survival (Zeh *et al.*, 2002), and a recent estimate of age of sexual maturity and the maximum age (George *et al.*, 1999). This information was included in the analyses of this chapter as data-based informative priors (see the section on model parameters and prior distributions).

Population dynamics model

The underlying population model is a simplified age- and sex- structured Leslie matrix (Leslie, 1945, 1948) projected as:

$$\underline{n}_{t+1} = (\mathbf{A}_t \underline{n}_t) - \underline{h}_t \quad (1.1)$$

where:

- \underline{n}_t is vector of population size in each age class at the start of year t (defined when calving and natural mortality occur);
- \mathbf{A}_t is the Leslie matrix for year t ; and
- \underline{h}_t is the vector of age-specific catches during year t .

The catches and birth rates are assumed to be equal for both males and females (i.e. the vectors above are divided equally by sex). The parameters that define the entries of the Leslie matrix are: 1) s_{juv} , the survival rate of immature whales (assumed identical for calves and juveniles); 2) a_T , the last age with survival rate s_{juv} ; 3) s_a , the survival rate of mature whales; 4) a_m , the age at sexual maturity (the last age class with zero fecundity); 5) f_{max} , the maximum fecundity rate; and 6) a_{max} , the maximum age, after which survival becomes zero. Fecundity is

assumed identical for all mature animals, and is calculated as the number of female calves per mature female. Recruitment to the fishery is assumed to be knife-edged and to occur at age 1, and the catch is distributed uniformly over all recruited age-classes (i.e., uniform selectivity across recruited age-classes).

The projections are initialized from a stable age distribution for the population in the year prior to that with the first catch (e.g. 1977 or 1847) based on the values for the parameters sampled from the prior. This population vector is then projected one year forward without catch, and the population vector re-scaled so that the 1+ population size in the year with the first catch equals that generated from the prior for 1+ abundance for that year.

Density dependence is assumed to affect fecundity according to:

$$f_t = f_0 + (f_{\max} - f_0) \left[1 - \left(\frac{N_t^{1+}}{K^{1+}} \right)^z \right] \quad (1.2)$$

where:

- f_t is the fecundity during year t ;
- f_{\max} is the maximum fecundity (in the limit of zero population size);
- N_t^{1+} is the (1+) population size at the start of year t ;
- K^{1+} is the pre-exploitation (1+) population size;
- z is the shape parameter; and
- f_0 is the fecundity at carrying capacity.

Given values for the life-history parameters in the model, the value for f_0 is determined from the characteristic equation of the Leslie matrix given equilibrium conditions:

$$f_0 = \frac{1 - S_a}{S_{juv}^{(a_T)} S_a^{(a_m - a_T)} [1 - S_a^{(a_{\max} - a_m - 1)}]} \quad (1.3)$$

Model parameters and prior distributions

Table 1.4 lists the parameters and their priors used in three ‘reference’ scenarios. The population trajectory is modeled in three ways: 1) a density-dependent model initialized in 1848 (abbreviated: ‘1848 DD’); 2) a density-dependent model initialized in 1978 (‘1978 DD’); and 3) a density-independent model initialized in 1978 (‘1978 NON DD’). The six life-history parameters of the Leslie model are included in each of three models, but the remaining parameters differ among models. The ‘1848 DD’ model includes a parameter for the population size in 1993, N_{1993}^{1+} , and one for the maximum sustainable yield level, $MSYL^{1+}$. The ‘1978 DD’ model also includes the parameter $MSYL^{1+}$ but instead of placing a prior on N_{1993}^{1+} places one on N_{1978}^{1+} . This model also includes an additional (explicit) prior on the carrying capacity, K^{1+} . The ‘1978 NON DD’ model includes priors on N_{1978}^{1+} and the maximum population growth rate in the absence of density dependence, r . For this model, fecundity and population growth rates apply only to the specified period, and where the distinction is appropriate, they are referred to as f and r . However, when methods are consistent across models, these rates are referred to as f_{\max} and r_{\max} for the sake of simplicity.

‘Data-based’ prior distributions are assigned to adult survival rate and the age at sexual maturity, and the maximum age of the Leslie matrix is determined from the results of recent research on ageing. The informative prior for s_a (Table 1.4) approximates the Bayesian posterior calculated for this parameter based on a mark-recapture analysis of photoidentification data (Zeh *et al.* 2002). Information on age-at-maturity is taken from a study by George *et al.* (1999) that estimated ages of caught animals based on the chemistry of eye lenses. Those authors fit a growth curve to these ages from known lengths, and combine this relationship with

previous data on length at sexual maturity to provide an estimate of the age-at-maturity. The aging results also estimated some animals were older than the previously accepted maximum age, and this result is supported by the recent recovery of traditional whaling tools in five whales (reported in the same study). The maximum age in the Leslie matrix is therefore set to 200 to reflect this information.

Punt and Butterworth (1999) note that placing a prior on S_{juv} would be an instance of Borel's Paradox (i.e. effectively placing two priors on the same parameter) due to the functional relationships among the life-history parameters in an age-structured population model; instead the value of S_{juv} is solved for analytically in this study by rearranging the characteristic equation of the Leslie matrix given the values for the remaining five parameters and λ , the dominant eigenvalue of the Leslie matrix (i.e. $r_{\max} + 1$) (Breiwick *et al.*, 1984):

$$S_{juv} = \left[\frac{\lambda^{(a_m+1)} - S_a \lambda^{(a_m)}}{S_a^{(a_m-a_T)} f_{\max} \{1 - (S_a / \lambda)^{(a_{\max}-a_m-1)}\}} \right]^{\frac{1}{a_T}} \quad (1.4)$$

The value for S_{juv} is forced to be less than that of S_a . If necessary, values for f_{\max} and S_a are re-sampled (see below), until this condition is met, or 1000 re-samples occur. If this maximum is reached, a new value for r_{\max} is re-sampled, and the process repeated until an acceptable sample from the prior occurs.

The priors for the remaining life history parameters are also based on available information. The prior distribution for f_{\max} is based on an assumed range of a 2.5- to 4-year calving interval for large baleen whales (IWC, 1998). Note that f_{\max} is specified in the standard Leslie matrix formulation as female calves per female per year (i.e., a fecundity rate of 0.125 implies a female calving interval of 8 years, and

therefore a total calving interval of 4 years, assuming an equal sex ratio of calves). The age of transition from immature to adult survival is assigned a discrete uniform prior over the interval 0 to 8 years.

Output quantities

Posterior distributions are calculated for several output quantities that are functions of the parameters in Table 1.4. The maximum sustainable yield rate ($MSYR^{1+}$) is calculated as $\lambda - 1$ based on the f_t value associated with $MSYL^{1+}$. Maximum sustainable yield (MSY^{1+}) is calculated as the product, $MSYR^{1+} MSYL^{1+} K^{1+}$. Current replacement yield (RY^{1+}) is calculated as the difference between the number of 1+ animals in 2002 (prior to the removal of catches in that year) and the number of such animals at the end of 2001. The quantity Q_1^{1+} , designed to meet the intent of aboriginal whaling management objectives (Wade and Givens 1997), is also calculated. This quantity has the property that the proportion of net production allocated to recovery increases at higher levels of stock depletion³. Specifically:

$$Q_1^{1+} = \min(MSYR^{1+} * N_{2002}^{1+}, 0.9MSY^{1+}) \quad (1.5)$$

The post-model-pre-data distribution is reported for the parameters. This distribution arises after conditioning the specified priors on the model (i.e. by eliminating combinations of parameters for which the juvenile survival rate implied by equation (1.4) exceeds the adult survival rate drawn from the joint prior distribution). Likewise, post-model-pre-data distributions for output quantities are calculated as the distributions for these quantities in the sampled joint prior space.

Parameter estimation

³ This definition applies to a population above some minimum level, P_{min} (assumed here to be $0.1K^{1+}$), below which catches are set to zero.

The Sampling-Importance-Resample (SIR) algorithm (Rubin, 1988; Smith and Gelfand, 1992; Wade, 2002a) is used to generate samples of parameter vectors (and output quantities of interest) from the posterior distribution. This algorithm involves randomly sampling a large number of parameter vectors θ_i (draws) from the prior distribution. A population trajectory is then calculated for each vector of parameter values, and this trajectory is used to determine the likelihood of the data for each random draw. 10,000 draws (which form the numerical representation of the posterior distribution) are then selected by sampling (with replacement) from the initial samples from the prior, with probability proportional to the likelihood. Following Punt and Butterworth (1999) and Raftery *et al.* (1995), the SIR algorithm is considered to have converged if the number of unique parameter vectors in the sample from the posterior is fairly high (>5,000) and if the most frequently re-sampled parameter vector did not occur in the posterior sample more than ten times.

The total negative log-likelihood of a model trajectory, given a vector of parameters and the data, consists of contributions from four data sources: 1) the estimate of abundance for 1993; 2) the estimates of abundance for the remaining years; 3) the proportion of calves in the population; and 4) the proportion of mature animals in the population. The abundance estimates are assumed to be indices of the 1+ component of the population. The scientific surveys at Pt. Barrow are assumed to have occurred after the aboriginal catch, and the likelihood function is calculated accordingly (i.e. catches are removed before calculating the likelihood of the data for a given year). Model-predicted proportions are calculated over the period 1985 to 1994, as the actual stage proportions are based on data for these years.

The estimate of abundance for 1993 is assumed to be independent of the remaining estimates (Punt and Butterworth, 1999), and to have normally as

opposed to log-normally distributed sampling error. The contribution of the abundance estimates to the negative of the log-likelihood function is:

$$L_1 \propto 0.5 \frac{(\hat{N}_{1993} - 8293)^2}{626^2} \quad L_2 \propto 0.5 \sum_{t_1} \sum_{t_2} (\ln \hat{N}_{t_1} - \ln N_{t_1}^{obs})^T \Sigma_{t_1, t_2}^{-1} (\ln \hat{N}_{t_2} - \ln N_{t_2}^{obs}) \quad (1.6)$$

where:

N_t^{obs} is the estimate of abundance for year t ;

\hat{N}_t is the model estimate of 1+ abundance for year t ; and

Σ is the variance-covariance matrix for the logarithms of the estimates of abundance (excluding 1993).

The estimates of abundance (Table 1.1) are based on combining the data from visual counts at Point Barrow, Alaska, and estimates of the proportion of animals which passed within visual range based on acoustic data. Equation (1.6) accounts for the correlation among the non-1993 estimates of abundance that arises because the proportion within visual range is treated as a random effect when constructing the estimates of abundance (Zeh and Punt, 2005).

The contribution of the proportion data to the likelihood function follows Punt (2006), i.e. given the bootstrapping approach adopted to calculate the length-frequency distributions from which the proportion data were calculated (Koski *et al.*, 2006), it was reasonable to assume that the estimates are normally distributed:

$$L_3 \propto \frac{1}{2(\sigma_{p_c})^2} (p_c - p_c^{obs})^2 \quad L_4 \propto \frac{1}{2(\sigma_{p_m})^2} (p_m - p_m^{obs})^2 \quad (1.7)$$

where:

p_c^{obs} is the observed fraction of the population that consisted of calves between 1985 and 1994;

- σ_{p_c} is the standard deviation of p_c^{obs} ;
- p_c is the model-estimate of the fraction of the population that consisted of calves between 1985 and 1994;
- p_m^{obs} is the observed fraction of the population that consisted of mature animals between 1985 and 1994;
- σ_{p_m} is the standard deviation of p_m^{obs} ; and
- p_m is the model-estimate of the fraction of the population that consisted of mature animals between 1985 and 1994.

Model comparison

The three models considered in this chapter are compared using Bayes factors (Kass and Raftery, 1995). The Bayes factor is calculated as the probability of observing the data given one hypothesis (model) divided by the probability of observing the same data given an alternative hypothesis, i.e.:

$$B_{12} = \frac{pr(D|H_1)}{pr(D|H_2)} \quad (1.8)$$

In the context of model comparison, the hypotheses represent competing models and the Bayes factor is used as the evidence provided by the data in favour of one model over another. Although equation (1.8) has the form of a likelihood ratio, if there are unknown parameters in either of the competing models, the probability densities must be found by integrating, as opposed to maximizing, over the parameter space. Therefore, for a given a model, the probability of the data is:

$$pr(D|H) = \int pr(D|H, \theta) pr(\theta|H) d\theta \quad (1.9)$$

This integration is based on the sample from the prior using the equation:

$$\hat{pr}(D|H) = \frac{1}{n_1} \sum_{i=1}^{n_1} pr(D|\underline{\theta}_i, H) \quad (1.10)$$

where $\underline{\theta}_i$ is the i^{th} (of n_1) samples from the prior distribution.

Model averaging

Model uncertainty is accounted for by calculating the posterior probability of each model conditioned on the data and the priors, and then combining results across models as a weighted average of the posterior densities for a quantity of interest (Kass and Raftery, 1995). Hoeting *et al.* (1999) provide a convenient method of calculating the posterior probability of model H_k (where, $k=1,2, \dots, K$ models are being considered) based on Bayes' theorem:

$$pr(H_k|D) = \frac{pr(D|H_k)pr(H_k)}{\sum_{i=1}^K pr(D|H_i)pr(H_i)} \quad (1.11)$$

where $pr(H_k)$ is the prior probability that H_k is the true model and $pr(D|H_k)$ is the estimate of the probability of the data (equation (1.10)). All of the probabilities are conditional on the set of models being considered (Hoeting *et al.*, 1999). For f_{max} , r_{max} and quantities related to carrying capacity, only the two models incorporating density dependence could be used to derive model averages. Under the set of models considered, these posterior model probabilities were used to determine model-averaged posterior probability distributions for the model outputs, θ :

$$pr(\theta|D) = \sum_{i=k}^K pr(\theta|H_k, D)pr(H_k|D) \quad (1.12)$$

In the context of the SIR algorithm used here, Bayesian model averaging was accomplished by selecting a number of random draws from the posterior for each

model and combining them to form a model-averaged posterior. This number was determined by the posterior probability for each model. All models were considered equally probable *a priori* (i.e. objective ignorance regarding the true model), so the posterior probability of a given model was determined using the values from equation (1.10) normalized to sum to one over models.

1.3 RESULTS

Fits to the data

Fig. 1.1 shows the fits to the abundance estimates for three reference scenarios and the Bayesian model average. All three models provide a relatively good fit to these data, although the abundance estimate for 2001 falls well above the upper 90% credibility limit from the ‘1848 DD’ model. This occurs because the median of the posterior distribution for 1+ population size in 2002 for this model (9,496) is smaller than the posterior medians for the other two models (Fig. 1.2, Table 1.5).

The abundance estimates indicate the population has been increasing steadily over 1978-2001, and the data on adult survival, age of sexual maturity, and the stage-proportion data (excluding 1985) are all relatively consistent with this increase (i.e., the inclusion of those data in the analysis does not lead to the model being unable to mimic the abundance data).

Backwards to 1848 (density dependent model): 1848 DD

The upper left panel of Fig. 1.1 shows the posterior median time-trajectory of 1+ population size along with its 90% credibility interval for the reference scenario. The population size is estimated to have declined dramatically during the 1800s, being reduced to approximately half of its pre-exploitation level within five years of the start of the commercial fishery, and 10% of this size by the early 1900s. However, the population recovered steadily thereafter. The 90% credibility interval for the post-model-pre-data distribution for K^{1+} is [8,000-30,000], with lower

values favoured (Fig. 1.2). In contrast, the 90% credibility interval for the posterior distribution of K^{1+} is [9,000, 14,000] (Table 1.5) indicating that the data update the prior distribution substantially. The 2002 population size is estimated to be above 50% of K^{1+} , and there is a high probability of it being above $MSYL^{1+}$ (Fig. 1.3, Table 1.5). The posterior distribution for replacement yield in 2002 has a mode around 200, with a lower 5th percentile of 61 (Fig. 1.4, Table 1.5). In contrast, the lower 5th percentile for Q_1^{1+} is 99 (Fig. 1.4, Table 1.5). These two quantities differ because the current population size is estimated to be larger than $MSYL^{1+}$, and approaching K^{1+} . Therefore, density dependence has slowed population growth and RY^{1+} has decreased. This is the same situation which led to the use of Q_1^{1+} as a more appropriate measure of sustainable catch (to achieve IWC management goals) for the eastern North Pacific stock of gray whales (Wade, 2002a).

The constraints imposed by the relationships among the life-history parameters constrained r_{\max} to be less than about 0.07 (Fig. 1.3), although they also reduced the (prior) probability of values of r_{\max} larger than 0.06. The posterior for r_{\max} assigns most support to values larger than 0.03 (posterior median 0.041, Table 1.6). The posterior distributions for adult and immature survival favour higher values than implied by the prior distributions for these parameters, and that for a_m values lower than implied by the prior distributions (Figs 1.5 and 1.6). The post-model-pre-data distribution for fecundity favours higher values (~ 0.20 , Fig. 1.6), but the posterior median is 0.171, or a calving interval of approximately 3 years (Table 1.6). The results in Figs 1.3, 1.5 and 1.6 show that the data are clearly capable of updating the prior distributions for the life-history parameters.

The results for this model are not particularly sensitive to changing the prior distributions for S_a and a_m and to ignoring the proportion data (Table 1.5). The most noteworthy feature of these sensitivity tests are the changes to the catch-

related outputs (RY^{1+} , Q_1^{1+} and MSY^{1+}). Results are not shown in Table 1.5 for the case in which the 1985 stage-proportion data are included in the analyses due to computational difficulties in achieving convergence. Preliminary analyses including these data indicated, however, that they are inconsistent with what is known about bowhead life history and the time series of abundance estimates. This inconsistency was the cause of the inability to achieve convergence.

Forwards from 1978 (density dependent model): 1978 DD

The posterior for K^{1+} from this analysis is much more uncertain than that from the ‘1848 DD’ analysis (Fig. 1.2). This is because this analysis ignores the information contained in the 1848-1977 catch record, and because the abundance estimates show no evidence for a reduction in trend (which would be expected as the population approaches carrying capacity) and, unlike the ‘1848 DD’ model, the ‘1978 DD’ model does not make the assumption that the population size in 1848 was K^{1+} . Therefore, although this model confirms that the population is increasing (Fig. 1.1), it infers that the population is currently at a much lower fraction of its (current) carrying capacity than the ‘1848 DD’ model (Fig. 1.3; Table 1.5). The posterior for N_{2002}^{1+}/K^{1+} is strongly influenced by the prior distribution assumed for K^{1+} given the inability of the data to place an upper bound on K^{1+} . This is clearly evident from the results of the sensitivity test in which the upper limit of the prior for K^{1+} is increased from 30,000 to 100,000. The results for this sensitivity test imply an increase to the median of the posterior for K^{1+} of 165% and a reduction to the posterior median for N_{2002}^{1+}/K^{1+} of 62% (Table 1.5) as the upper bound for K^{1+} is increased by 233%.

Given that there is little independent information on which to base a prior distribution for K^{1+} , the choice of the prior for K^{1+} is essentially arbitrary, and it

should be recognized that this ‘forwards’ analysis consequently does not provide robust estimates of quantities related to K^{1+} (such as N_{2002}^{1+}/K^{1+}). However, Punt and Butterworth (1999) note that some key management-related quantities (e.g., RY^{1+} and Q_1^{1+}) are relatively insensitive to the prior assumed for K^{1+} , so this approach still has some value. The implicit (post-model-pre-data) distribution for RY^{1+} favours (is skewed towards) values less than 200 (Fig. 1.4). Despite this, the posterior median is 324, with a lower 5th percentile of 147 (Table 1.5). A similar result is evident for Q_1^{1+} , with low values favoured by the post-model-pre-data distribution, but higher values supported by the data (median 295, lower 5th percentile 160)(Fig. 1.4, Table 1.5). In essence, the joint prior distribution for the parameter values, conditioned on the population dynamics model, is not neutral (non-informative) with respect to these catch-related quantities, but the data are influential enough to move the posterior distribution away from the mode of the prior distribution.

The posterior distributions for r_{\max} , $MSYR^{1+}$ and the life history parameters for this model are generally similar to those for the ‘1848 DD’ model. The most noteworthy difference between the posterior distributions for the ‘1978 DD’ and ‘1848 DD’ models in Table 1.6 relates to the posterior median for S_{juv} which is larger for the ‘1978 DD’ model. As was the case for the ‘1848 DD’ model, there is again little sensitivity to changing the priors for S_a and for a_m , and ignoring the stage-proportion data (Tables 1.5 and 1.6). In contrast, inclusion of the 1985 survey data when calculating the stage-proportion data has a large impact on the results. Specifically, K^{1+} and $MSYR^{1+}$ are estimated to be lower, and N_{2002}^{1+}/K^{1+} higher, with the population estimated to be above $MSYL^{1+}$ with almost 100% probability. r_{\max} is estimated to be higher (as the population is estimated to be closer to K^{1+}

and therefore experiencing a growth rate much lower than r_{\max}). The estimates of the catch-related quantities are considerably lower for this scenario (e.g. the posterior median and the lower 5% percentile for RY^{1+} are 166 and 44 respectively).

Forwards from 1978 (density independent model): 1978 NON DD

The posterior median for RY^{1+} for this model ranges from 310 to 414 across the scenarios (166-217 for the lower 5th percentile for RY^{1+}). The posterior for r for the ‘1978 NON DD’ model is centred on lower values than those for the other two models (Fig. 1.3). This is to be expected because the r for the ‘1978 NON DD’ model’ pertains to the current rate of increase rather than the increase rate in the limit of zero population size. The inclusion in the analyses of the 1985 stage-proportion data is again very influential. For example, the posterior distribution for fecundity for the ‘include 1985 proportion data’ sensitivity test does not overlap with that for the reference scenario.

Model comparison and Bayesian model averages

Bayes factors based on pair-wise comparisons of models range from 1.10 to 1.51, and indicate that there is no evidence for selecting one model over another (Table 1.7). Rather, these Bayes factors imply that the best approach to summarizing the state of the B-C-B bowhead stock is to consider all three models, e.g. through Bayesian model averaging. Average likelihoods of draws from the initial sample range from 0.522 (‘1848 DD’) to 0.789 (‘1978 NON DD’). The two models that involve forward projection from recent abundance (‘1978 DD’ and ‘1978 NON DD’) have slightly higher average likelihoods, and hence posterior model probabilities, than the model which started the population projection in 1848, although differences are not large (Table 1.8).

The time-trajectory of 1+ population size (medians and 90% credibility intervals) from the Bayesian model-averaged posterior provides, as expected, a good fit to the abundance estimates (Fig. 1.1, lower panel). The fit to the estimate of abundance for 2001 for the model-averaged posterior is not quite as good as for the ‘1978 DD’ and ‘1978 NON DD’ models because of the impact of including the ‘1848 DD’ model in the average. The model-averaged posterior distribution for K^{1+} (Fig. 1.2) has a mode close to that of the posterior median for the ‘1848 DD’ model, and a long tail caused by the uncertainty associated from the ‘1978 DD’ model. The model-average posterior for N_{2002}^{1+}/K^{1+} (Fig. 1.3) is wide, but less so than that for the ‘1978 DD’ model. The model-averaged posterior for RY^{1+} is slightly irregular because it consists of the combination of a bimodal posterior (for the ‘1848 DD’ model) and a symmetric posterior (for the ‘1978 DD’ model). In contrast to the model-average posterior for RY^{1+} , that for Q_1^{1+} is quite symmetric (Fig. 1.4, Table 1.5).

The ‘1978 NON DD’ model estimates only the recent fecundity and rate of increase for the population, whereas the two density-dependent models estimate the maximum fecundity and rate of increase. Therefore, the posterior for f_{\max} and r_{\max} is averaged across the two models with density dependence only. Maximum fecundity and population growth rate are relatively consistent across these two models and have a median of 0.171 and 4.3% respectively (Fig. 1.6 and Fig. 1.3, Table 1.6). Likewise, both adult and juvenile survival rates are consistent across models, with a median for adult survival of 0.990 and for juvenile survival of 0.932 (Fig. 1.5, Table 1.6).

It is straightforward to calculate model-averaged posterior probability distributions given different prior probability distributions for the models. For example, the models based on starting the projections in 1978 could be assigned

probabilities of 0.25 and that which starts the population projections in 1848, a prior probability of 0.5 to indicate, for example, that the assumption that carrying capacity has not changed over the last 150 years is equally as likely as some shift in the equilibrium population size during this time. Alternatively, the models with density-dependence could be assigned prior probabilities of 0.25 each, and that which ignores density-dependence a prior probability of 0.5.

1.4 DISCUSSION

The three models have shown a good concurrence. However, use of the historic catch record leads to lower estimates of RY^{1+} because the analysis estimates the population to be close to carrying capacity and so the growth rate is reduced compared to that at low population size. However, there is no (visual) evidence in the abundance estimates for a reduction in trend. It therefore appears that it is the combination of the magnitude of the historical catches and the values for the biological parameters that determines the estimate of the carrying capacity. One implication of starting the population projection in 1848 is that the model under-predicts the 2001 estimate of abundance (probably because if the population is approaching carrying capacity, a near-linear growth in population size could not still be occurring). The ‘1848 DD’ model is assigned less weight than the ‘1978 DD’ and ‘1989 NON DD’ models using Bayes factor, but the discrepancy between the predictions of the ‘1848 DD’ model and the data remains sufficiently small that the ‘1848 DD’ model cannot be rejected.

RY^{1+} is the catch that will keep a population at its current size. This quantity is less useful as the basis for management advice for the B-C-B bowhead stock now that at least some of the analyses suggest the recruited population may be approaching K^{1+} . Obviously RY^{1+} will be zero if the population stops increasing because it reaches carrying capacity. Q_1^{1+} is therefore a more appropriate catch-

related quantity to examine because it does not become zero at carrying capacity. Furthermore, this quantity represents a catch level that has been argued to meet the requirements of aboriginal subsistence management (Wade and Givens, 1997; Wade, 2002a). The fact that there is no evidence to select one model over the others, and not all models result in similar estimates of catch quantities is a reason why model uncertainty is important to include when conducting assessments of marine renewable resources. The lower 5th percentile of the Bayesian model-averaged posterior of Q_1^{1+} is 155 whales, and represents our best estimate of the catch level that would meet the intent of aboriginal whaling management objectives, taking into account both parameter, and model uncertainty (to the extent that model uncertainty can be captured by the three models considered in this study). It should be noted, in light of recent discussions regarding stock structure (IWC, 2005b), that the results presented here are based on the assumption that the B-C-B population of bowhead whales is composed of a single stock.

The actual aboriginal catch quotas are driven by need, and have averaged 36 whales per year from 1978-2002. There appears to be little effect on population size due to this catch level. In fact, during this timeframe the B-C-B stock of bowhead whales is estimated to have more than doubled. Another way of putting an average take of 36 whales per year into perspective is to examine the annual net production over the last 25 years. If it is assumed that the population was increasing at a constant 3.5% per year (the median rate estimated from the density independent Leslie matrix analysis), the population increased by about 175 whales per year in 1978 (population size ~5,000), about 260 whales per year in 1990 (population size ~ 7,500), and about 350 whales per year in 2002 (population size ~ 10,000). Given that the population has increased from about 5,000 whales to about 10,000 whales, in this timeframe an average kill of 36 whales per year represents an annual catch rate between 0.35-0.70% of the total population size.

What is known about B-C-B bowhead life-history vital rates (survival, fecundity, etc.) appears consistent with the available data on trends in abundance and the proportion of the population in three stages (calves, immature and mature). Overall, the results support a value of r_{\max} of between 0.03-0.05, a range often assumed for cetaceans, particularly species with delayed sexual maturity and a longer than 2-year calving interval (Reilly and Barlow 1986, Wade 1998, Wade 2002b). In light of the reproductive life-history of this species, the results make clear that the observed population growth rates can only be supported by extremely high survival rates, as already suggested by the estimates of adult survival (0.990) in Zeh *et al.* (2002), and the observations of exceptionally old individuals (George *et al.*, 1999).

The analyses of this chapter are based on the same types of data that were available for the 1998 assessment of this stock. Several other sources of data exist. For example, Schweder and Ianelli (2000) examine whether the data on the age-composition of the 1973-93. Catches are consistent with the abundance and proportion data. Punt (2006) shows that it is possible to reconcile the abundance, proportion, length-frequency and age-composition data within a Bayesian framework. It would be straightforward conceptually (but perhaps computationally challenging) to use the approach outlined in this chapter to compare models that utilize these additional data sources. Bringing in those additional data, as well as doing a full model comparison of a variety of models, was beyond the scope of this chapter, but we agree this would be important future work, particularly in light of the methods now developed in Punt (2006).

This study represents the first attempt to quantify model uncertainty when conducting assessments of the B-C-B bowhead stock. The analyses consider three alternative models and take model uncertainty into account by weighting alternative models based on their posterior model probabilities and by calculating a Bayesian

model-averaged posterior. The only previous attempt to consider model uncertainty when conducting assessments of whale stocks was by Wade (2002a), who compared models for the Eastern North Pacific stock of gray whales with and without additional variance about the abundance estimates. In that case, one model received almost all the weight making model-averaging redundant. In contrast, in this study all three models were assigned non-negligible weight and led to different estimates of quantities of interest (e.g., carrying capacity and related measures). The Bayesian model-averaged posterior distribution clearly represents our best efforts to incorporate all levels of uncertainty in the estimates of these quantities.

Table 1.1

Estimates, CVs (actually the standard errors of the log abundance estimates, which are approximately equal) and the correlation matrix for the indices of abundance for the Bering-Chukchi-Beaufort Seas stock of bowhead whales. Source: Zeh and Punt (2005).

Year	Estimate	CV	Correlation matrix										
1978	4,765	0.305	1.000										
1980	3,885	0.343	0.118	1.000									
1981	4,467	0.273	0.056	0.050	1.000								
1982	7,395	0.281	0.094	0.084	0.035	1.000							
1983	6,573	0.345	0.117	0.104	0.049	0.084	1.000						
1985	5,762	0.253	0.070	0.062	0.020	0.078	0.062	1.000					
1986	8,917	0.215	0.072	0.064	0.017	0.092	0.064	0.113	1.000				
1987	5,298	0.327	0.124	0.110	0.052	0.088	0.110	0.065	0.067	1.000			
1988	6,928	0.120	0.028	0.025	0.013	0.017	0.024	0.009	0.007	0.026	1.000		
1993	8,167	0.071	0.001	0.001	0.001	0.000	0.001	-0.001	-0.002	0.001	0.000	1.000	
2001	10,545	0.128	0.008	0.007	0.005	0.001	0.007	-0.004	-0.008	0.008	0.003	0.000	1.000

Table 1.2

The proportion of observed calves (p_c^{obs}) and mature (p_m^{obs}) animals with associated standard errors, over the years 1985 to 1994. Proportions are given based on including and ignoring the anomalous 1985 data set, as well as those used in the previous assessment. Source: IWC (1999) and Koski *et al.* (2006).

Scenario	p_c^{obs}	σp_c	p_m^{obs}	σp_m
Exclude 1985	0.0580	0.0062	0.4366	0.0106
Include 1985	0.0309	0.0034	0.4160	0.0096
1998 assessment	0.052	0.0164	0.411	0.0286

Table 1.3

Catches of Bering-Chukchi-Beaufort Seas bowhead whales, 1848–2002. Values in parenthesis are the catches used by Punt and Butterworth (1999) in the 1998 assessment where these catches differ from those used in the present analyses.

Year	Total Kill	Year	Total Kill	Year	Total Kill	Year	Total Kill
1848	18	1887	240	1926	35	1965	14
1849	573	1888	160	1927	14	1966	24
1850	2067	1889	127	1928	30	1967	12
1851	898	1890	136	1929	30	1968	27
1852	2709	1891	284	1930	17	1969	32
1853	807	1892	346	1931	32	1970	48
1854	166	1893	180	1932	27	1971	25
1855	2	1894	234	1933	21	1972	44
1856	0	1895	117	1934	21	1973	51
1857	78	1896	118	1935	15	1974	42
1858	461	1897	130	1936	24	1975	32
1859	372	1898	309	1937	53	1976	74
1860	221	1899	234	1938	36	1977	72
1861	306	1900	148	1939	18	1978	15
1862	157	1901	55	1940	20	1979	20
1863	303	1902	162	1941	38	1980	32
1864	434	1903	116	1942	26	1981	26
1865	590	1904	86	1943	14	1982	14
1866	554	1905	105	1944	8	1983	16
1867	599	1906	69	1945	23	1984	16
1868	516	1907	96	1946	20	1985	14
1869	382	1908	123	1947	21	1986	22
1870	637	1909	61	1948	8	1987	29
1871	138	1910	37	1949	11	1988	28
1872	200	1911	48	1950	23	1989	25
1873	147	1912	39	1951	23	1990	41
1874	95	1913	23	1952	11	1991	47
1875	200	1914	61	1953	41	1992	46
1876	76	1915	23	1954	9	1993	51
1877	270	1916	23	1955	36	1994	39 (38)
1878	80	1917	35	1956	11	1995	56 (57)
1879	266	1918	27	1957	5	1996	42 (45)
1880	480	1919	33	1958	5	1997	62
1881	435	1920	33	1959	2	1998	51
1882	242	1921	9	1960	33	1999	47
1883	42	1922	39	1961	17	2000	42
1884	160	1923	12	1962	20	2001	67
1885	377	1924	41	1963	15	2002	44
1886	168	1925	53	1964	24		

Table 1.4

Prior distributions used for the reference scenarios for each model type. Dashes (-) represent prior distributions that are equal to those from the model in the column to the left (e.g. the 1978 NON DD model had the same prior on the size of the 1+ population in 1978 as the 1978 DD model). 'N/A' represents a prior that was not applicable to a certain model or models. Fecundity is defined as female calves per mature female. For the 1978 NON DD model r_{\max} and f_{\max} are not maxima, but constant values. Results from the reference scenario were used in all figures and when performing the Bayesian model averaging. The reference scenarios are based on the stage-proportion data set that ignores the data for 1985.

Parameter		Model Type		
		1848 DD	1978 DD	1978 NON DD
S_a	adult survival	N(0.990, 0.02), truncated at 0.940 and 0.995 ^a	-	-
f_{\max}	maximum fecundity	U[0.125, 0.200] ^b	-	-
a_T	age-at-transition to adult survival	U[0, 8] ^c	-	-
a_m	age-at-sexual maturity	N(20.0, 3.0) truncated at 13.0 and 26.0 ^d	-	-
$r_{\max} = \lambda - 1$	intrinsic population growth rate	U[0.005, 0.075] ^e	-	-
N_{1978}^{1+}	1+ population size in 1978	N/A	U[3000, 9000] ^f	-
N_{1993}^{1+}	1+ population size in 1993	N(7800, 1200) ^g	N/A	N/A
K^{1+}	1+ carrying capacity	N/A	U[8000, 30000] ^h	N/A
$MSYL^{1+}$	$MSYL$ in terms of the 1+ component	U[0.40, 0.80] ⁱ	-	-

Table 1.4 continued

- a – based on the posterior distribution for adult survival rate obtained by Zeh *et al.* (2002).
- b – the prior for the maximum number of calves (of both sexes) per mature female selected by the Scientific Committee of the International Whaling Commission was $U[0.25, 0.4]$ (IWC, 1995). This is the corresponding prior given fecundity has been defined here as female calves per mature female per year.
- c – selected by the Scientific Committee of the International Whaling Commission (IWC, 1995) although there is little information on the value of this parameter (Givens *et al.*, 1995).
- d – based on a best estimate of 20 years and a lower confidence for the age-at-maturity of 14 years (IWC, 1995).
- e – preliminary trials indicated there was no posterior probability outside this range, which was confirmed in the final analyses. This range was therefore selected to improve the efficiency of the numerical integration while not affecting the results.
- f – selected to encompass a plausible range of values for 1+ population size in 1978.
- g – selected by the Scientific Committee of the International Whaling Commission (IWC, 1995) based on the prior distribution assumed for the Bayes empirical Bayes estimate of abundance (Raftery and Zeh, 1991).
- h – based on the range selected by the Scientific Committee of the International Whaling Commission (IWC, 1995).
- i – selected to encompass the range of values commonly assumed when conducting assessments of cetacean populations.

Table 1.5

Posterior medians (5th, 95th percentiles) for eight management-related quantities. This table includes results for all models and scenarios, and Bayesian model-averaged results for the reference scenarios. Model averages for all quantities, except RY and N_{2002}^{1+} (i.e., the two quantities not dependent on K), could only be based on the 1848 and 1978 DD models. Where noted, additional scenarios are based on the reference scenarios in Table 1.4, but changed to examine the sensitivity of the results to different proportion data sets and less informative priors. The uniform prior for S_a is $U[0.940, 0.995]$, and that for a_m is Discrete $U[13, 26]$. Unique draws and the maximum number of times an individual draw is included in the final SIR resample are listed in rightmost two columns.

	N_{2002}^{1+}	K^{1+}	RY^{1+}	Q_1^{1+}	N_{2002}^{1+} / K^{1+}
1848 DD					
Reference scenario	9496 (8750, 10180)	10960 (9190, 13950)	171 (61, 233)	228 (149, 296)	0.888 (0.647, 0.985)
No proportion data	9380 (8652, 10070)	10980 (9245, 14710)	160 (63, 222)	216 (132, 283)	0.874 (0.602, 0.983)
Uniform priors on S_a and a_m	9488 (8782, 10180)	10580 (9112, 13610)	159 (59, 229)	241 (158, 301)	0.921 (0.674, 0.986)
1978 DD					
Reference scenario	10670 (9042, 12410)	20510 (11010, 29120)	324 (147, 501)	295 (160, 439)	0.530 (0.356, 0.925)
No proportion data	10410 (8740, 12380)	20350 (10600, 29070)	297 (119, 519)	270 (136, 457)	0.524 (0.348, 0.930)
Include 1985 proportion data	9294 (7780, 10720)	13510 (8110, 28890)	166 (44, 280)	193 (74, 289)	0.705 (0.299, 0.990)
Uniform priors on S_a and a_m	10820 (9124, 12600)	19870 (10750, 29090)	336 (138, 525)	309 (168, 459)	0.554 (0.362, 0.947)
$U[8000, 100000]$ prior on K	10830 (9160, 12690)	54430 (14990, 95500)	363 (188, 566)	267 (136, 426)	0.201 (0.110, 0.710)
1978 NON DD					
Reference scenario	10740 (9130, 12700)	N/A	366 (204, 588)	N/A	N/A
No proportion data	11020 (9199, 13130)	N/A	402 (206, 654)	N/A	N/A
Include 1985 proportion data	10280 (8796, 12090)	N/A	310 (166, 511)	N/A	N/A
Uniform priors on S_a and a_m	11110 (9265, 13050)	N/A	414 (217, 644)	N/A	N/A
Bayesian model average	10276 (8907, 12406)	13854 (9466, 28475)	297 (92, 539)	257 (155, 412)	0.720 (0.372, 0.980)

Table 1.5 continued

	$MSYL^{1+}$	MSY^{1+}	$MSYR^{1+}$	Unique draws	max sampled draw (n)
1848 DD					
Reference scenario	0.734 (0.639 , 0.793)	253 (166 , 329)	0.033 (0.019 , 0.048)	7399	9
No proportion data	0.703 (0.555 , 0.791)	240 (147 , 315)	0.032 (0.017 , 0.047)	9456	4
Uniform priors on S_a and a_m	0.740 (0.644 , 0.795)	267 (176 , 335)	0.035 (0.020 , 0.049)	6527	9
1978 DD					
Reference scenario	0.651 (0.573 , 0.779)	368 (206 , 599)	0.030 (0.018 , 0.043)	8992	5
No proportion data	0.661 (0.528 , 0.785)	341 (169 , 644)	0.029 (0.016 , 0.043)	9519	3
Include 1985 proportion data	0.607 (0.459 , 0.784)	225 (89 , 333)	0.025 (0.009 , 0.048)	8034	8
Uniform priors on S_a and a_m	0.654 (0.570 , 0.781)	379 (215 , 624)	0.032 (0.019 , 0.045)	7493	9
U[8000, 100000] prior on K	0.606 (0.523 , 0.777)	712 (282 , 1737)	0.025 (0.014 , 0.038)	7554	7
1978 NON DD					
Reference scenario	N/A	N/A	N/A	9262	5
No proportion data	N/A	N/A	N/A	9809	3
Include 1985 proportion data	N/A	N/A	N/A	6264	9
Uniform priors on S_a and a_m	N/A	N/A	N/A	8374	6
Bayesian model average	0.695 (0.581 , 0.788)	302 (180 , 555)	0.031 (0.018 , 0.046)		

Table 1.6

Posterior medians (5th, 95th percentiles) for eight life history-related quantities for all models and scenarios, and the Bayesian model-averaged results for the reference scenario. For the 1978 NON DD model, fecundity and population growth rate only apply to the specified period, and are referred to here as f and r . The model-averaged results for these rates are based only on the two models with density dependence, and refer to f_{\max} and r_{\max} .

	S_a	S_j	f_{\max} (or f)	r_{\max} (or r)
1848 DD				
Reference scenario	0.989 (0.977, 0.995)	0.926 (0.718, 0.980)	0.171 (0.135, 0.198)	0.041 (0.024, 0.059)
No proportion data	0.991 (0.978, 0.995)	0.957 (0.739, 0.990)	0.181 (0.136, 0.199)	0.046 (0.024, 0.062)
Uniform priors on S_a and ASM	0.988 (0.972, 0.994)	0.943 (0.786, 0.981)	0.174 (0.138, 0.198)	0.044 (0.026, 0.059)
1978 DD				
Reference scenario	0.991 (0.979, 0.995)	0.945 (0.765, 0.988)	0.171 (0.133, 0.198)	0.045 (0.025, 0.063)
No proportion data	0.991 (0.978, 0.995)	0.957 (0.753, 0.990)	0.181 (0.136, 0.199)	0.045 (0.025, 0.062)
Include 1985 proportion data	0.993 (0.979, 0.995)	0.981 (0.750, 0.993)	0.182 (0.126, 0.199)	0.049 (0.027, 0.065)
Uniform priors on S_a and a_m	0.989 (0.974, 0.995)	0.956 (0.816, 0.988)	0.171 (0.135, 0.198)	0.047 (0.027, 0.064)
U[8000, 100000] prior on K	0.990 (0.978, 0.995)	0.940 (0.752, 0.987)	0.162 (0.130, 0.197)	0.042 (0.025, 0.061)
1978 NON DD				
Reference scenario	0.990 (0.977, 0.995)	0.923 (0.717, 0.978)	0.141 (0.127, 0.163)	0.035 (0.022, 0.049)
No proportion data	0.989 (0.975, 0.995)	0.936 (0.684, 0.985)	0.173 (0.132, 0.198)	0.038 (0.022, 0.053)
Include 1985 proportion data	0.989 (0.976, 0.995)	0.823 (0.477, 0.899)	0.197 (0.191, 0.200)	0.031 (0.018, 0.044)
Uniform priors on S_a and ASM	0.988 (0.973, 0.995)	0.943 (0.783, 0.982)	0.144 (0.128, 0.166)	0.038 (0.023, 0.052)
Bayesian model average	0.990 (0.978, 0.995)	0.932 (0.733, 0.984)	0.171 (0.133, 0.198)	0.043 (0.025, 0.062)

Table 1.6 continued

	a_m	a_T	p_c	p_m
1848 DD				
Reference scenario	16 (14 , 21)	5 (1 , 9)	0.055 (0.046 , 0.065)	0.436 (0.418 , 0.453)
No proportion data	17 (13 , 23)	5 (1 , 9)	N/A	N/A
Uniform priors on S_a and ASM	15 (13 , 20)	5 (1 , 9)	0.056 (0.048 , 0.065)	0.437 (0.420 , 0.454)
1978 DD				
Reference scenario	16 (13 , 21)	5 (1 , 9)	0.057 (0.048 , 0.066)	0.435 (0.418 , 0.452)
No proportion data	17 (13 , 23)	5 (1 , 9)	N/A	N/A
Include 1985 proportion data	20 (13 , 25)	5 (1 , 9)	0.034 (0.025 , 0.040)	0.424 (0.394 , 0.663)
Uniform priors on S_a and a_m	15 (13 , 19)	5 (1 , 9)	0.058 (0.050 , 0.066)	0.436 (0.419 , 0.453)
U[8000, 100000] prior on K	16 (14 , 21)	5 (1 , 9)	0.057 (0.049 , 0.067)	0.435 (0.417 , 0.453)
1978 NON DD				
Reference scenario	17 (14 , 22)	5 (1 , 9)	0.060 (0.054 , 0.068)	0.434 (0.417 , 0.450)
No proportion data	19 (14 , 23)	5 (1 , 9)	N/A	N/A
Include 1985 proportion data	18 (14 , 22)	4 (1 , 9)	0.081 (0.080 , 0.084)	0.424 (0.414 , 0.437)
Uniform priors on S_a and ASM	15 (13 , 21)	5 (1 , 9)	0.061 (0.054 , 0.069)	0.435 (0.418 , 0.452)
Bayesian model average	16 (14 , 21)	5 (1 , 9)	0.058 (0.049 , 0.067)	0.435 (0.417 , 0.452)

Table 1.7

Bayes factors for comparison of paired models. Evidence categories are modifications of the original categories of Jefferys (1961), as presented by Kass and Raftery (1995) and used by Wade (2002a) in an assessment of the Eastern North Pacific gray whales: >150 is decisive evidence, 12-150 is strong evidence, 3-12 is positive evidence, and 1-3 is not worth more than a bare mention. All comparisons are based on the results of the reference scenarios.

Models	Bayes factor	Evidence for the first model
1978 DD <i>vs</i> 1848 DD	1.37	Not worth more than a bare mention.
1978 NON DD <i>vs</i> 1848 DD	1.51	""
1978 NON DD <i>vs</i> 1978 DD	1.10	""

Table 1.8

The average likelihood (Eq. 10) and posterior model probabilities used in the Bayesian model averaging. It was only possible to consider the two DD models for averaging quantities related to carrying capacity. All models were considered equally likely *a priori*.

Model	Average Likelihood	Posterior model probabilities for models considered	
		All three	1848 DD and 1978 DD
1848 DD	0.522	0.258	0.422
1978 DD	0.715	0.353	0.578
1978 NON DD	0.789	0.389	N/A

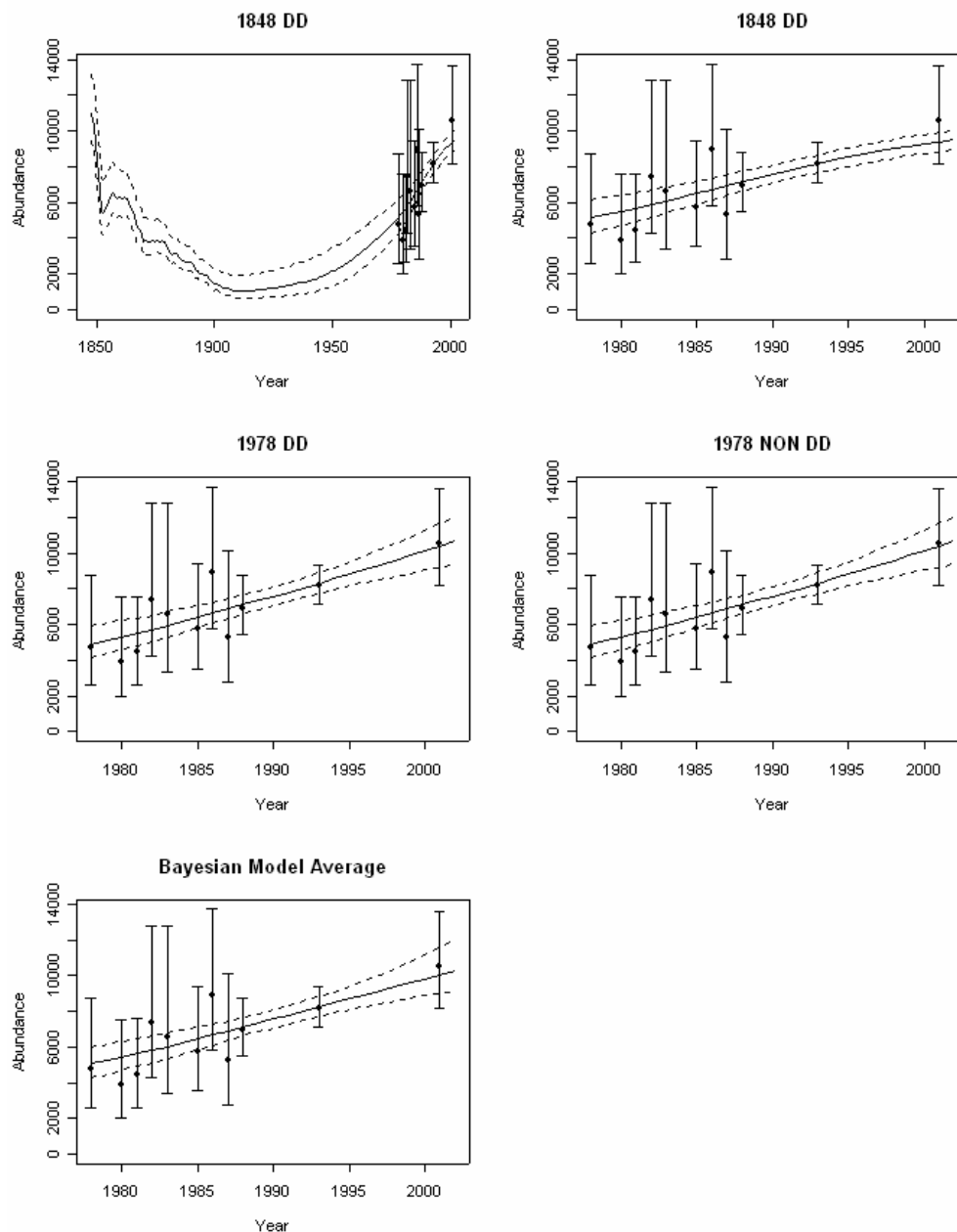


Figure 1.1: Time trajectories (medians and 90% credibility intervals) for 1+ population size for the three reference scenarios and the Bayesian model average. The two uppermost plots are for the 1848 DD model, showing the entire trajectory from 1848, and only the recent trajectory from 1978 for comparison. Error bars represent 95% CI's, and are assumed to be log-normally distributed for all abundance estimates except 1993 (second to last), which is assumed to be normally distributed.

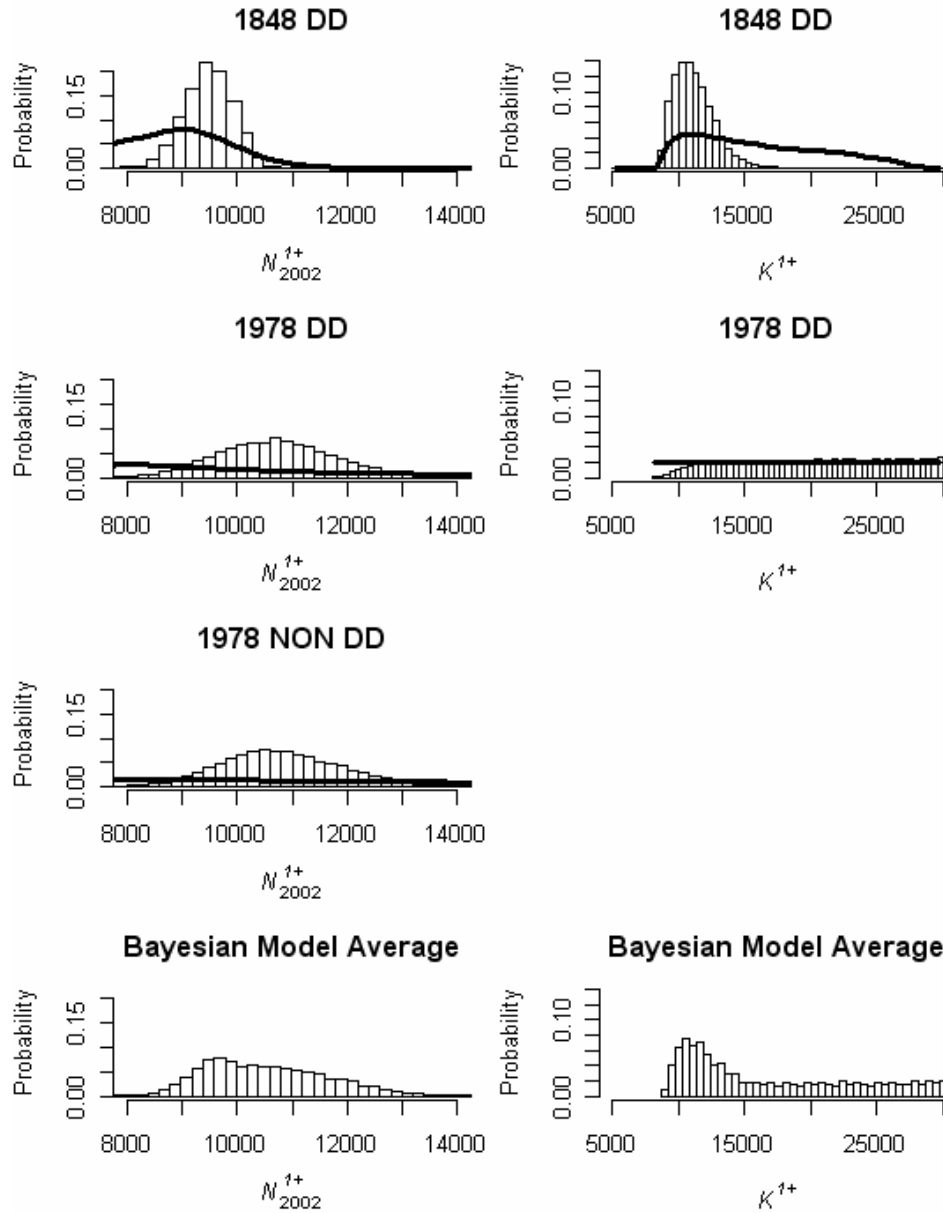


Figure 1.2: and Posterior (vertical bars) and post-model-pre-data (solid lines) distributions for 1+ population size in 2002, N_{2002}^{1+} (left panels) and 1+ carrying capacity, K^{1+} (right panels). Results are shown for only two of the three reference scenarios for K^{1+} , and the Bayesian model average for K^{1+} is based on the results of these two scenarios only.

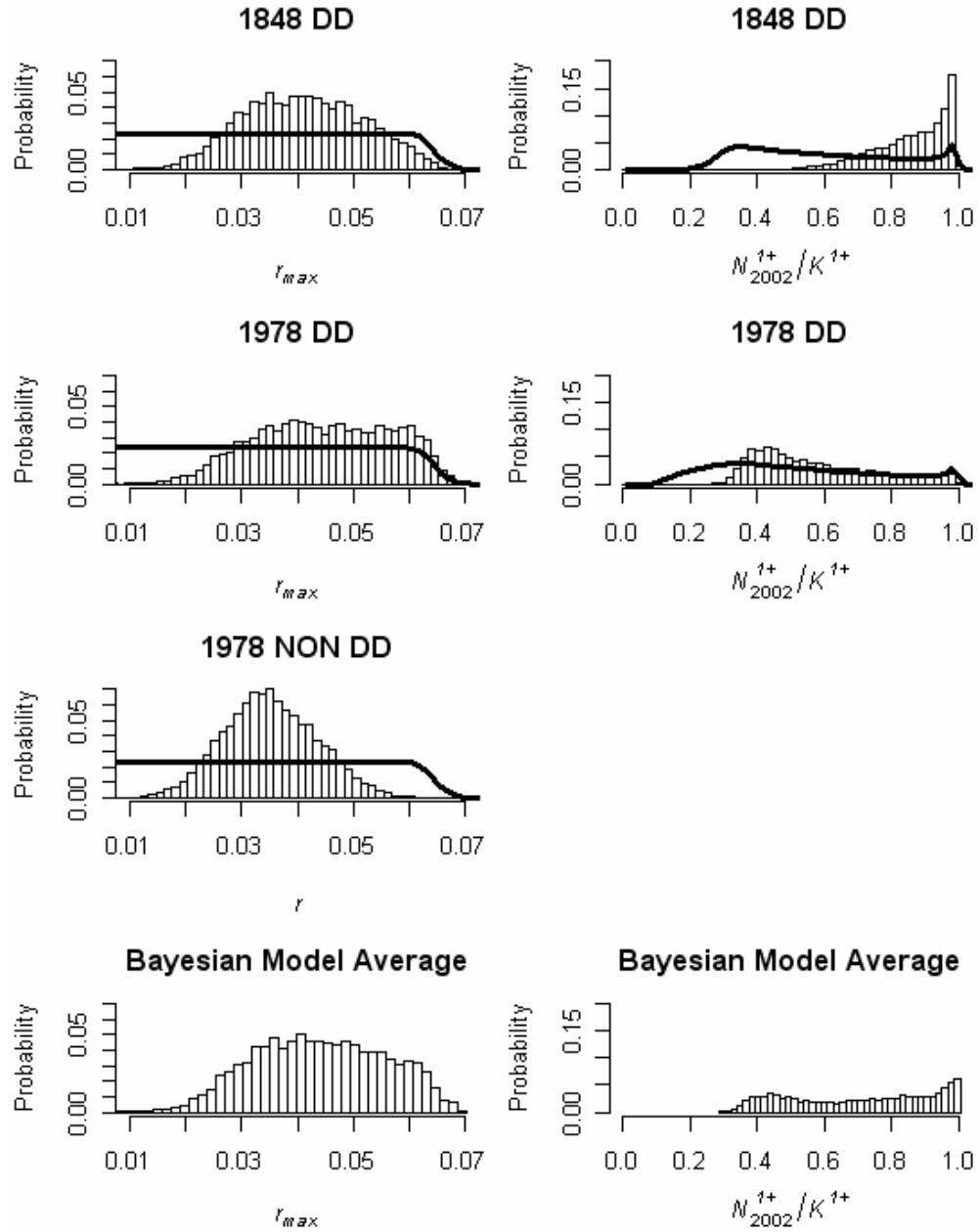


Figure 1.3: Posterior (vertical bars) and post-model-pre-data (solid lines) distributions for intrinsic population growth rate, r_{max} and r (left panels) and recent depletion in terms of the 1+ component, N_{2002}^{1+}/K^{1+} (right panels). The Bayesian model average is based only on the two models with density dependence.

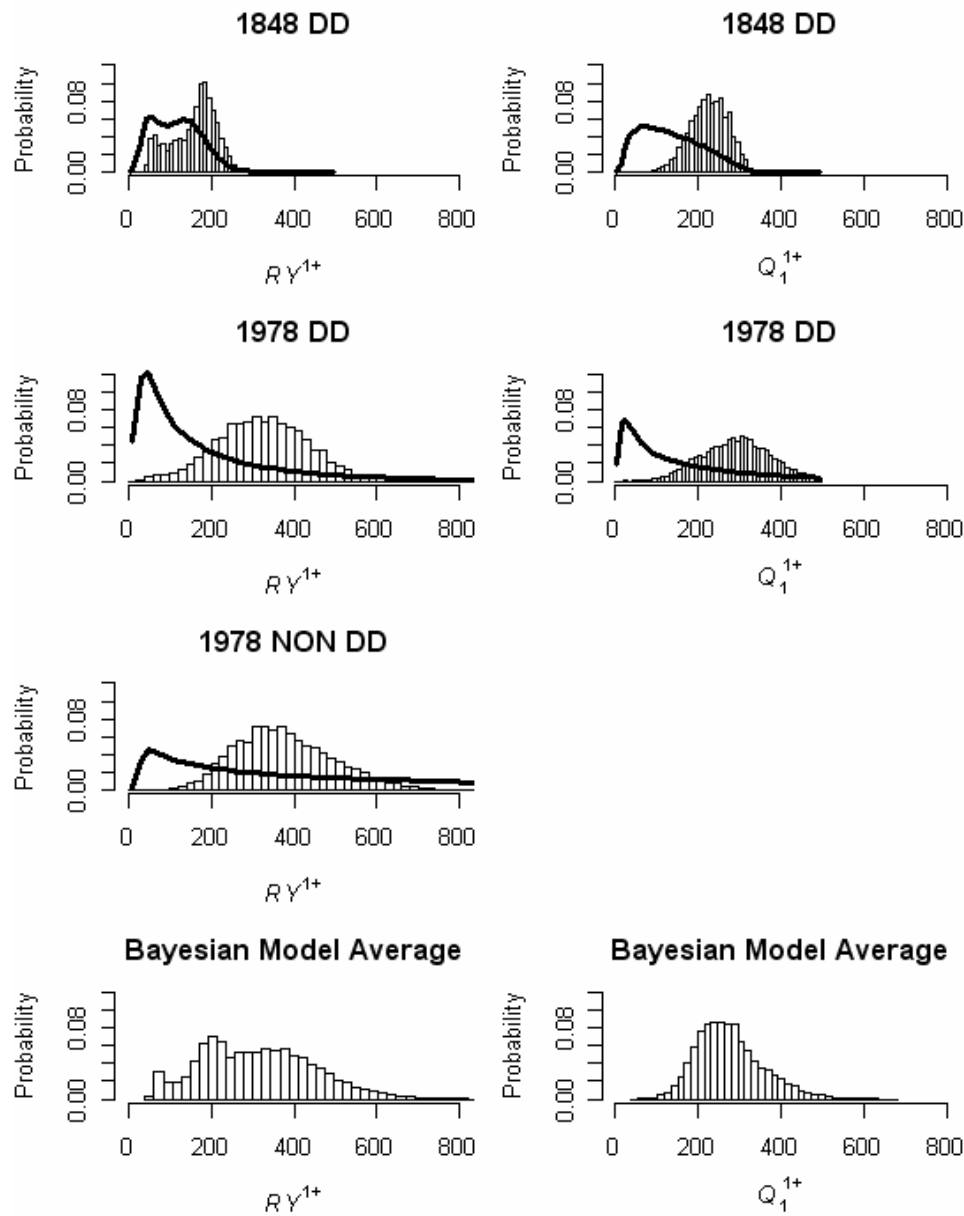


Figure 1.4: Posterior (vertical bars) and post-model-pre-data (solid lines) distributions for replacement yield, RY^{1+} (left panels) and Q_1^{1+} (right panels). Results are shown for only two of the three reference scenarios for Q_1^{1+} , and the Bayesian model average for Q_1^{1+} is based on the results of these two scenarios only.

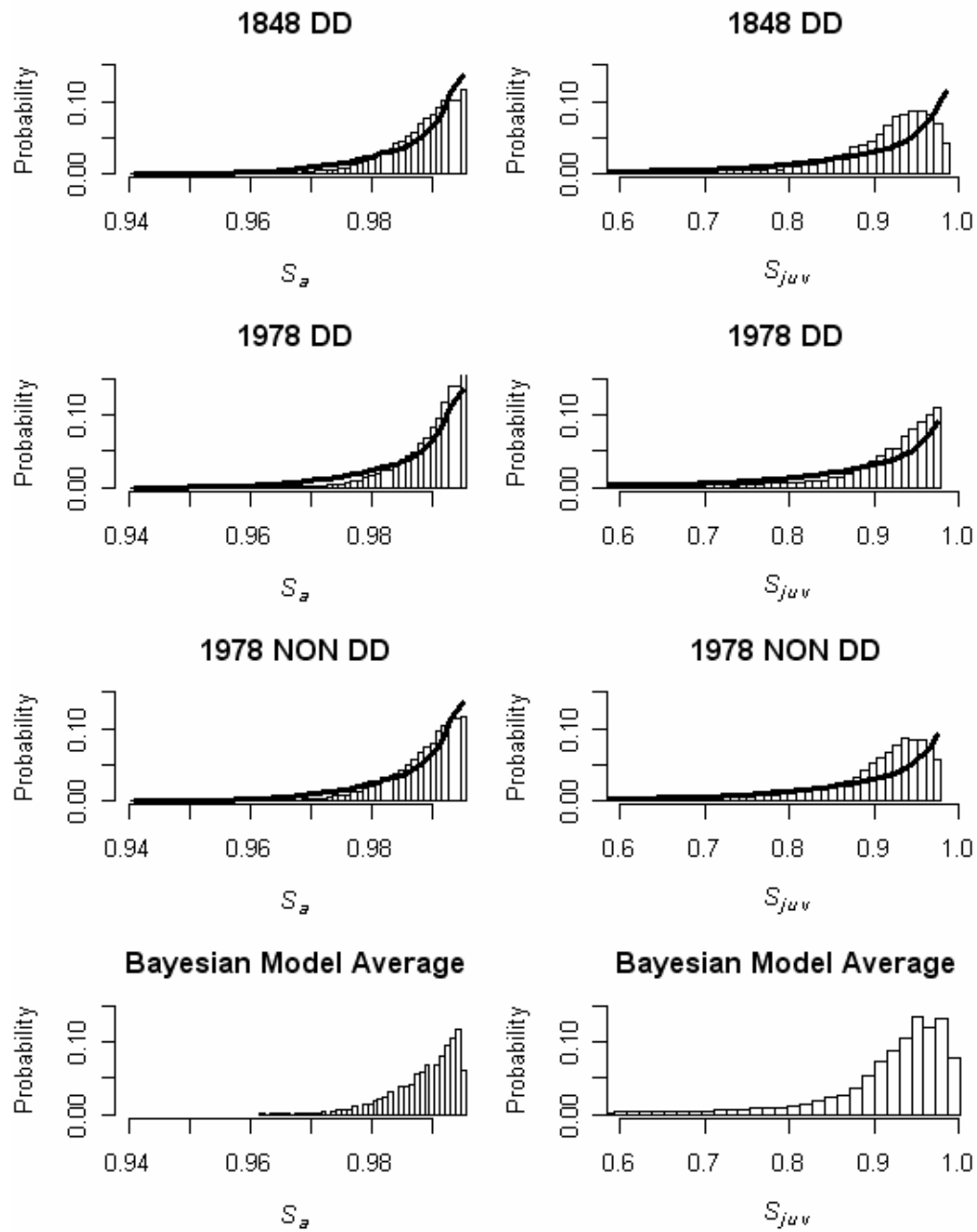


Figure 1.5: Posterior (vertical bars) and post-model-pre-data (solid lines) distributions for adult survival rate, S_a (left panels) and calf and juvenile survival rate, S_{juv} (right panels). Results are shown for the three reference scenarios and for the Bayesian model average.

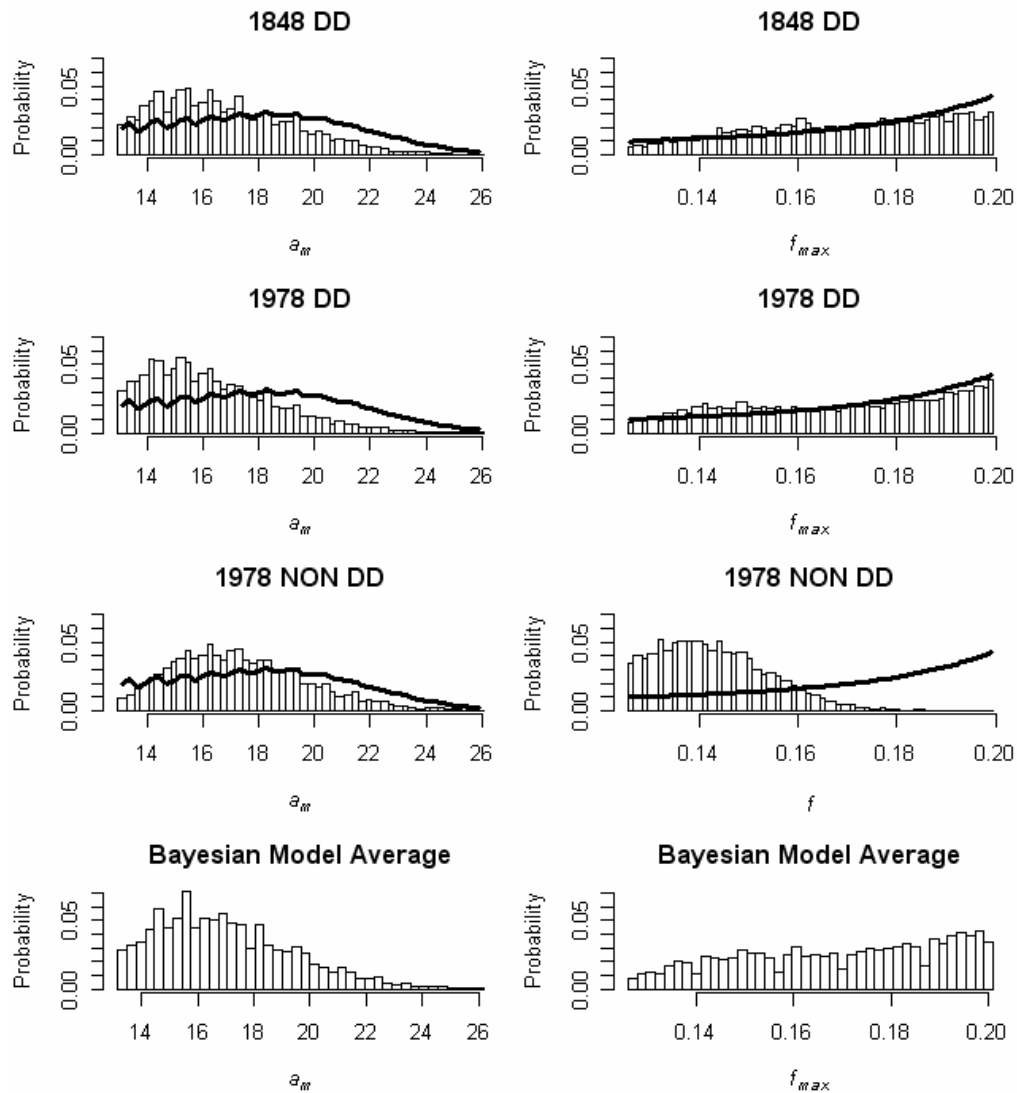


Figure 1.6: Posterior (vertical bars) and post-model-pre-data (solid lines) distributions for the age-at-maturity, a_m (left panels) and fecundity, f_{max} or f (right panels). Results are shown for the three reference scenarios and for the Bayesian model average, which is based only on the two models with density dependence for fecundity.

Chapter 2:

Constructing a coherent joint prior while respecting biological realism: sensitivity of marine mammal stock assessments to recent methods⁴

ABSTRACT

Bayesian estimation methods, employing the Sampling-Importance-Resampling algorithm, are currently used to perform stock assessments for several stocks of marine mammals, including the Bering-Chukchi-Beaufort Seas stock of bowhead whales (*Balaena mysticetus*) and walrus (*Odobenus rosmarus rosmarus*) off Greenland. However, due to the functional relationships among parameters in deterministic age-structured population dynamics models, placing explicit priors on each life-history parameter in addition to the population growth rate parameter results in an incoherent prior distribution (i.e., two different priors on the same parameter). One solution to constructing a coherent joint prior is to solve for juvenile survival analytically, given values generated from the prior distributions for the remaining parameters in the model (including population growth rate). However, certain combinations of parameter values result in solutions for juvenile survival that are larger than adult survival, which is biologically implausible. Therefore, in order to respect biological realism, parameter values must be rejected for some or all of the remaining parameters. This study investigates several alternative resampling schemes for obtaining feasible solutions for juvenile survival. The sensitivity of assessment results is investigated for data-rich (bowhead) and data-poor (walrus) scenarios. The results based on limited data are especially sensitive to the choice of alternative resampling scheme.

⁴ This chapter was published in ICES Journal of Marine Science:

Brandon, J.R., Breiwick, J.M., Punt, A.E. and Wade, P.R. 2007. Constructing a coherent joint prior while respecting biological realism: application to marine mammal stock assessments. *ICES J. Mar. Sci.* 64: 1085-1100.

2.1 INTRODUCTION

Bayesian methods utilizing age-structured population dynamics models (PDMs) have formed the basis for recent stock assessments of several marine mammal populations, including those for Atlantic walrus, *Odobenus rosmarus rosmarus*, in Greenland (Witting and Born, 2005); the Bering-Chukchi-Beaufort Seas (B-C-B) stock of bowhead whales, *Balaena mysticetus* (Brandon and Wade, 2006; Punt, 2006); British grey seals, *Halichoerus grypus* (Buckland *et al.*, 2004; Thomas *et al.*, 2005); the Eastern North Pacific stock of gray whales, *Eschrichtius robustus* (Wade, 2002a; Witting, 2003; Punt *et al.*, 2004); New Zealand sea lions, *Phocarctos hookeri* (Breen *et al.*, 2003; Breen and Kim, 2006); Northeastern offshore spotted and Eastern spinner dolphins, *Stenella attenuata* and *Stenella longirostris orientalis* (Wade, 1994; Alvarez-Flores, 2002; Hoyle and Maunder, 2004); and Steller sea lions, *Eumetopias jubatus* (Fay, 2004; Fay and Punt, 2006; Winship and Trites, 2006). These assessments use Bayesian estimation methods to calculate posterior probability distributions for model parameters and management-related quantities, and to form the basis for risk analyses to evaluate the implications of potential management actions (Hilborn and Walters, 1992; Ellison, 1996). A brief review is provided below of the developments required for commonly used Bayesian methods when applied to age-structured PDMs for marine mammal stock assessments: the necessity of constructing a coherent joint prior distribution; one solution to this challenge that has been employed in several recent stock assessments; and, finally, the importance of investigating the sensitivity of those results to alternative schemes for constructing a coherent joint prior while respecting biological realism.

Adopting similar approaches to Butterworth *et al.* (1987) and Nakamura *et al.* (1989), Givens (1993), Raftery *et al.* (1995) and Givens *et al.* (1995) developed and influenced the application of Bayesian methods to age-structured PDMs for marine mammal stock assessment. The method developed, ‘Bayesian synthesis,’ was a

departure from the ‘standard Bayesian’ approach in that it allowed explicit prior distributions on *both* model inputs *and* model outputs. During the development of these methods, it was shown that Bayesian synthesis (as originally proposed) is inappropriate on theoretical grounds (Wolpert, 1995; Schweder and Hjort, 1996). Givens and Raftery (1997) clarified the concerns by distinguishing two problems: 1) Borel’s paradox (i.e., conditioning on an ill-defined distribution, resulting in this case from explicit priors on both model input and outputs, and ultimately leading to posterior distributions that are not invariant to model parameterization); and, 2) prior incoherence (i.e., the partly implicit presence of two different prior distributions on the same parameter). Bravington (1996) provides illustrative examples of these issues using a simple model of exponential population growth. Borel’s paradox does not apply to a standard Bayesian analysis. However, the issue of prior incoherence is a potential concern for all Bayesian stock assessments. As Givens and Roback (1999) stress, simultaneous and competing prior distributions for the same parameters are common occurrences in Bayesian modeling.

A coherent joint prior must be constructed before inference may be drawn from Bayesian analyses. Punt and Hilborn (1997) advise that care needs to be taken in order to avoid *implicitly* specifying contradictory priors for the same parameter in Bayesian stock assessments. Those authors demonstrate how explicit prior distributions on both unexploited biomass and current depletion (the ratio of current biomass to unexploited biomass) lead to incoherent priors. Such subtleties become more insidious as PDMs (and the functional relationships among parameters therein) increase in complexity. Fortunately, several approaches have been developed to address this issue in the context of age-structured PDMs. However, we postpone consideration of alternative methods at this stage, and instead return to this during the concluding discussion. The sensitivity analyses presented here focus on only one approach for constructing a coherent joint prior while respecting biological realism: the most commonly employed during recent assessments of

large whales as outlined below (e.g., Punt and Butterworth, 1999; 2002; Wade, 2002a; Punt *et al.* 2004; Punt, 2006; Brandon and Wade, 2006).

Due to the functional relationships among life-history parameters and the parameter related to population growth rate in a deterministic age-structured PDM (Lotka, 1907; Euler, 1970), placing priors on life history parameters in addition to a prior on the population growth rate would result in effectively placing two (most likely incoherent) priors on the same parameter. Therefore, one option for constructing a coherent joint prior is to place marginal priors on all but one of these parameters, and then solve for the value of the remaining parameter analytically. Given that many of the key model outputs (e.g. the current rate of increase) are almost directly proportional to the population growth rate, having an explicit prior on the latter parameter allows clear consideration of the impact of the priors on these model outputs. For example, placing a uniform prior on the population growth rate, to the extent that it is possible, essentially imposes a non-informative prior on some of the model outputs of interest to management.

There exists very little information on which to base a prior on juvenile survival S_{juv} for the majority of marine mammal populations. Therefore, it has been argued that placing an explicit prior on this life history parameter is unfounded for most species, and that a more appealing approach would be to simply impose some biological bounds instead of a full probability distribution (Butterworth, 1995). Additionally, a common assumption in age-structured PDMs for marine mammals, is that juvenile survival is less than adult survival S_a (e.g. IWC, 1995; Chivers, 1999; Witting and Born, 2005). This assumption is based on observed mammalian mortality patterns, which are typically U-shaped with age (Caughley, 1966; Barlow and Boveng, 1991). Thus, one solution to constructing a coherent joint prior is not to place an explicit prior on S_{juv} (Punt and Butterworth, 1996), but instead to solve for this parameter analytically given the values for the other life history parameters and the population growth rate (Breiwick *et al.*, 1984; Punt, 1999a). Furthermore,

the solution for S_{juv} is subject to the constraint, $S_{juv} < S_a$. This approach is appealing because it is a relatively simple, minimizes the influence of priors for life history parameters on key model outputs of interest to management, and results in a coherent joint prior distribution that also respects biological realism.

Sensitivity tests to alternative prior specifications should be conducted during the assessment processes to ensure that inference is robust. Effectively, the constraint $S_{juv} < S_a$ is an element of the joint prior on the life history parameters. However, certain combinations of otherwise reasonable parameter values result in solutions for S_{juv} which seem biologically implausible - if not completely impossible. Specifically, solutions for S_{juv} may result that are larger than those for adult survival (including values larger than 1.0; Figure 2.1). In these cases, it is possible to construct a coherent joint prior in more than one way. For example, in addition to completely ignoring that parameter space which violates the constraint, one could also resample values for different subsets of the life history parameters (effectively re-weighting the marginals) until a feasible solution for S_{juv} is attained. Both of these approaches satisfy the constraint $S_{juv} < S_a$ and both have been applied in marine mammal stock assessments. However, the question of sensitivity to such alternative resampling schemes has not been explored in any detail to date.

The Sampling-Importance-Resampling⁵ (SIR) algorithm for approximating Bayesian posterior distributions is employed in this chapter to generate realized prior distributions (also referred to as post-model-pre-data distributions). These distributions represent the actual (implicit) marginal prior distribution for each parameter after the explicit marginal prior distributions have been reconciled (via the model and resampling schemes) into a coherent joint probability distribution. It would be expected that the realized prior distribution will differ under alternative resampling schemes, but little is known about whether and how, if at all, the

⁵ ‘Resampling’ parameters values to find a feasible solution for S_{juv} should not be confused with the ‘resampling’ step in the SIR algorithm.

posteriors for the quantities of management interest are affected. The results of these interactions are potentially relevant to the calculation of management-related quantities and hence to the provision of management advice.

This chapter therefore explores the sensitivity of Bayesian assessments for marine mammal populations to several alternative resampling schemes used to construct a coherent joint prior, while simultaneously respecting biological realism. Results are provided for two populations that differ in terms of the amount of available data to illustrate the consequences in terms of quantities of management interest. The B-C-B bowhead stock is data-rich in that considerable information is available on abundance and trends in abundance (e.g. Zeh and Punt, 2005); in contrast the walrus off East Greenland (E-G) are data-poor, with only one (fairly imprecise) estimate of abundance on which to base assessments, and hence management advice (Born *et al.*, 1997; Witting and Born, 2005).

2.2. METHODS

To ease subsequent comparison, similar methods are used for the analyses of both the bowhead and walrus data. Hence, unless noted otherwise, the methods described here apply to both populations. Similarly, all parameters related to abundance or catch statistics are reported in terms of the 1+ component of the population (all age groups except calves), except for the trajectories of walrus population size which are plotted on the same scale as the prior for recent abundance (total population size including calves).

Available data

B-C-B bowhead whales

The PDMs utilise three sources of data: 1) abundance estimates from ice-based surveys at Pt. Barrow, Alaska between 1978 and 2001 (Table 2.1); 2) proportion calves/mature animals in the population from 1985 to 1994 (Table 2.2); and 3)

annual catches in individuals from 1848 to 2002 (see Punt, 2006 for this catch table).

E-G walrus

A single abundance estimate of 1000 exists for the E-G walrus population. This estimate is based on opportunistic and systematic observations (Born *et al.*, 1997; Witting and Born, 2005) and is assumed to relate to the total population (i.e., including young of the year). Following Witting and Born (2005), the coefficient of variation for this estimate is arbitrarily set to 0.35 to encompass the plausible range for abundance in 1995. This estimate forms the basis for the prior on abundance in 1995 (see below).

Born *et al.* (1997) report, or in some years estimate, catches of walrus off East Greenland during 1889-1999. These catches are treated as known (or estimated without error) in the model. No attempt has been made in this analysis to take into account the potential number of animals struck and lost, or the numbers landed and not reported in a given year (i.e., this corresponds to the “low” catch history analyzed by Witting and Born (2005)). The reader is referred to Born *et al.* (1997) and Witting and Born (2005) for the catch table and a detailed list of sources regarding these catches.

Prior to 1956, there is no information on the sex-ratio of the catch. In 1956, walrus off East Greenland were protected north of ca. 72°N, effectively ending the foreign fishery on this population. After 1956, the sex-ratio in the Greenlandic catch is highly skewed towards males (~90%; Born *et al.*, 1997). Therefore, we follow the methods of Witting and Born (2005), and assume an even sex ratio prior to 1956, and a 9:1 M:F sex-ratio thereafter.

Inflated catches

The implications of higher levels of catch on model outputs are examined by multiplying the catch history by five, for both B-C-B bowheads and E-G walrus.

Population dynamics model

The underlying PDM is an age- and sex- structured Leslie matrix (Leslie, 1945, 1948) projected as:

$$N_{t+1} = (A_t N_t) - C_t \quad (2.1)$$

where:

- N_t is the matrix of population size by sex and age class at the start of year t (defined when births and natural mortality occur);
- A_t is the Leslie matrix for year t ; and
- C_t is the matrix of sex- and age-specific catches during year t .

The parameters that define the entries of the Leslie matrix are: 1) S_{juv} , the survival rate of immature animals (assumed identical for calves and juveniles); 2) a_T , the last age with survival rate S_{juv} ; 3) S_a , the survival rate of mature animals; 4) a_m , the age at sexual maturity (the last age-class with zero fecundity, i.e., birth occurs at age a_m+1 years, the age at first parturition); 5) f_{max} , the maximum fecundity rate; and 6) a_{max} , the age after which survival becomes zero. Fecundity is assumed to be identical for all mature animals, and is calculated as the number of female calves per mature female. The sex ratio at birth is assumed to be 50:50 male:female. Recruitment to the fishery is assumed to be knife-edged and to occur at age 1, and the catch is hence distributed uniformly over all recruited age-classes.

Density dependence and initial conditions

Density dependence is assumed to affect fecundity according to a Pella-Tomlinson functional relationship based on the depletion of the 1+ component of the population (Pella and Tomlinson, 1969; Allen, 1976):

$$f_t = f_0 + (f_{\max} - f_0) \left[1 - \left(\frac{N_t^{1+}}{K^{1+}} \right)^z \right] \quad (2.2)$$

where:

- f_t is the fecundity during year t ;
- N_t^{1+} is the (1+) population size at the start of year t ;
- K^{1+} is the pre-exploitation (1+) population size;
- z is the Pella-Tomlinson shape parameter; and
- f_0 is the fecundity at carrying capacity.

Given values of the life-history parameters, the value for f_0 is determined from the characteristic equation of the Leslie matrix given equilibrium conditions (Breiwick *et al.*, 1984; Punt, 1999a):

$$f_0 = \frac{1 - S_a}{S_{\text{juv}}^{(a_T)} S_a^{(a_m - a_T)} [1 - S_a^{(a_{\max} - a_m - 1)}]} \quad (2.3)$$

The population projections are initialized from a stable age distribution at the start of the year before the first catch is removed, given values for the parameters sampled from the joint prior distribution.

Modeling approaches

B-C-B bowhead whales

The population trajectory is modeled in two ways: 1) a density-dependent model initialized in 1848 (abbreviated: ‘1848 Bkwd’), and; 2) a density-dependent model initialized in 1978 (‘1978 Fwd’). The six life-history parameters of the Leslie model are included in each model. However the approach used to estimate the equilibrium population size, or carrying capacity, K , differs between the two modelling approaches. The ‘1848 Bkwd’ model includes a parameter with associated prior for the population size in 1993, N_{1993} , and the ‘backwards’ method

(Butterworth and Punt, 1995, 1999) is used to back-calculate to the population size in 1848 (assumed to be equal to K). Instead of placing a prior on N_{1993} , the ‘1978 Fwd’ model involves placing a prior on the population size in 1978, N_{1978} , and projecting ‘forwards’ from that initial year (Wade, 2002a). This model includes an additional (explicit) prior on K . However, unlike the ‘1848 Bkwd’ model, it does not make the assumption that the catch history is known without error or that K has remained unchanged over the last 150 years. Both models include a parameter for the depletion at which MSY is achieved (referred to as the ‘maximum sustainable yield level’ or $MSYL$), and one for the maximum population growth rate in the limit of no density dependence, λ_{\max} (the dominant real eigenvalue of the Leslie matrix). These two modeling approaches have been used in recent assessments of the B-C-B bowhead whale (e.g., Brandon and Wade, 2006); it is therefore of interest to include them both in this analysis for the sake of comparison.

E-G walrus

The analyses for this population are based on a density-dependent model initialized in 1899 (corresponding to the model ‘1848 Bkwd’ for B-C-B bowheads). This model also includes a parameter for the population size in 1995, N_{1995} in addition to the six life-history parameters and $MSYL$. Note that the assumption that harvest selectivity is uniform above age 1 differs from the assumption that selectivity increases with age from age 0 to age 10 made by Witting and Born (2005). However, this difference is inconsequential for the analyses of this chapter.

Model parameters and prior distributions

Calves are defined as young of the year (i.e., age 0) and f_{\max} is specified in the standard Leslie matrix formulation as female calves per female per year (e.g., a fecundity rate of 0.125 implies a female calving interval of 8 years, and therefore a total calving interval of 4 years, assuming an equal sex ratio of calves).

B-C-B bowhead whales

‘Data-based’ prior distributions are assigned to adult survival rate and the age-at-maturity (Table 2.3), and the maximum age in the Leslie matrix is set to 200 years following the results of recent research on ageing (George *et al.*, 1999).

E-G walrus

The prior distributions for the life history parameters for walrus are based on various field and modeling studies (Table 2.4). The age at which survival changes from immature to adult is fixed, and set to age 3. The maximum age after which survival becomes zero is set to 60 years.

While we attempt to follow the methods of Witting and Born (2005) in most regards, this study treats juvenile survival rate differently. We choose to follow methods recently employed in stock assessments used by the IWC (e.g., Punt and Butterworth, 1999; Wade, 2002a), i.e. we place a uniform prior from 1.01 to 1.12 on λ_{\max} , and do not place a prior on juvenile survival. The details of this treatment are given in the following section (and apply to the bowhead analyses as well).

Alternative resampling schemes

The assumption that juvenile survival must be less than adult survival is followed here. However, as Punt and Butterworth (1999) note, placing an explicit prior on S_{juv} (in addition to priors on the remaining life history parameters, and a prior on the population growth rate) would result in an incoherent prior due to the functional relationships among the life-history parameters. Instead, the value of S_{juv} is calculated analytically by rearranging the characteristic equation of the Leslie matrix given the values for the remaining five life history parameters and λ_{\max} (Breiwick *et al.*, 1984):

$$S_{\text{juv}} = \left[\frac{\lambda_{\text{max}}^{(a_m+1)} - S_a \lambda_{\text{max}}^{(a_m)}}{S_a^{(a_m-a_T)} f_{\text{max}} \left[1 - (S_a / \lambda_{\text{max}})^{(a_{\text{max}}-a_m-1)} \right]} \right]^{\frac{1}{a_T}} \quad (2.4)$$

The resulting value for S_{juv} is forced to be less than that of S_a through one of the alternative resampling schemes below. These schemes involve resampling the following parameters from their prior distributions if $S_{\text{juv}} > S_a$:

- 1) f_{max} , S_a and a_m ;
- 2) f_{max} and S_a ; and,
- 3) No parameters (abbreviation ‘None’), meaning the current vector of life-history values is simply ignored and a completely new set of values is drawn from the marginals (see ‘Parameter Estimation’).

Values for the parameters are re-sampled until $S_{\text{juv}} < S_a$, or 1000 re-samples occur for resampling schemes 1) and 2). If this maximum is reached, a new value for λ_{max} is selected, and the process is repeated until an acceptable sample from the prior occurs.

Output quantities

Posterior distributions are calculated for several output quantities that are derived from the parameters in Tables 2.3 and 2.4. The maximum sustainable yield rate (*MSYR*) is calculated as $(\lambda_{\text{MSY}} - 1)$ based on the fecundity value, f_{MSY} , associated with *MSYL*. The quantity Q_1 , designed to meet the intent of IWC aboriginal whaling management objectives (Wade and Givens 1997), is also calculated. This quantity has the property that the proportion of net production allocated to recovery increases the more depleted a population is assessed to be⁶. Specifically:

$$Q_1 = \min(MSYR * N_t, 0.9MSY) \quad (2.5)$$

⁶ This definition applies to a population above some minimum level, P_{min} , (assumed here to be 0.1K), below which catches are set to zero.

where N_t is the population size in 2000 for walrus or 2002 for bowheads.

The realized prior is reported for the parameters. This distribution arises after conditioning the specified priors on the model and the resampling scheme by eliminating combinations of parameters for which: 1) the juvenile survival rate implied by Equation (2.4) exceeds the adult survival rate drawn from the joint prior distribution, and 2) population trajectories that go extinct. Likewise, post-model-pre-data distributions for output quantities are calculated as the distributions for these quantities in the sampled joint prior space.

Parameter estimation

The SIR algorithm (Rubin, 1988; Smith and Gelfand, 1992) is used to generate samples of parameter vectors (and output quantities of interest) from the posterior distribution. This algorithm involves randomly sampling a large number of parameter vectors from the prior distribution. A population trajectory is then calculated for each vector of parameter values, and this trajectory is used to determine the likelihood of the data for each random draw⁷. 10,000 draws (which form a numerical representation of the posterior distribution) are then selected by sampling (with replacement) from the initial samples from the prior, with probability proportional to the likelihood (i.e., the importance function is set equal to the joint prior, so the importance weight is the likelihood). Following Punt and Butterworth (1999) and Raftery *et al.* (1995), the SIR algorithm is considered to have converged if the number of unique parameter vectors in the sample from the posterior is fairly high (>5,000) and if the most frequently re-sampled parameter vector does not occur in the posterior sample more than ten times.

B-C-B bowhead whales

The total negative log-likelihood of a model trajectory, given a vector of parameters and the data, consists of contributions from four data sources: 1) the

⁷ The likelihood is 1 for the walrus case because there are no abundance data for this population except for that on which the prior for the abundance in 1995 is based.

estimate of abundance for 1993; 2) the estimates of abundance for the remaining years; 3) the proportion of calves in the population; and 4) the proportion of mature animals in the population. The abundance estimates are assumed to be indices of the 1+ component of the population. The scientific surveys at Pt. Barrow are assumed to have occurred after the aboriginal catch⁸, and the likelihood function is calculated accordingly (i.e. catches are removed before calculating the likelihood of the data for a given year). Model-predicted proportions are calculated over the period 1985 to 1994, as the actual stage proportions are based on data for these years.

The estimate of abundance for 1993 is assumed to be independent of the remaining estimates (Punt and Butterworth, 1999), and to have normally rather than log-normally distributed sampling error. The contribution of the abundance estimates to the negative of the log-likelihood function is (ignoring constants independent of model parameters):

$$L_1 = 0.5 \frac{(\hat{N}_{1993} - 8293)^2}{626^2} \quad L_2 = 0.5 \sum_{t_1} \sum_{t_2} (\ln \hat{N}_{t_1} - \ln N_{t_1}^{obs})^T \mathbf{V}_{t_1, t_2}^{-1} (\ln \hat{N}_{t_2} - \ln N_{t_2}^{obs}) \quad (2.6)$$

where:

N_t^{obs} is the survey estimate of abundance for year t ;

\hat{N}_t is the model estimate of 1+ abundance for year t ; and

\mathbf{V} is the variance-covariance matrix for the logarithms of the estimates of abundance (excluding 1993).

The estimates of abundance (Table 2.1) are based on combining the data from visual counts at Pt. Barrow, Alaska, and estimates of the proportion of animals that passed within visual range based on acoustic data. Equation (2.6) accounts for the

⁸ In reality, there are two seasonal (spring and fall) hunts each year, with the survey immediately following the spring hunt. However, as catches are a relatively small proportion of the total population size, the simplification made will hardly affect the quantitative results.

correlation among the non-1993 estimates of abundance that arises because the proportion within visual range is treated as a random effect when constructing the estimates of abundance (Zeh and Punt, 2005).

The contribution of the proportion data to the likelihood function follows Punt (2006), i.e. given the bootstrapping approach adopted to calculate the length-frequency distributions from which the proportion data were calculated (Koski *et al.*, 2006), it is reasonable to assume that the estimates are normally distributed, i.e. ignoring constants independent of model parameters:

$$L_3 = 0.5 \frac{(p_c - p_c^{\text{obs}})^2}{(\sigma_{p_c})^2} \quad L_4 = 0.5 \frac{(p_m - p_m^{\text{obs}})^2}{(\sigma_{p_m})^2} \quad (2.7)$$

where:

p_c^{obs} is the observed average fraction of the population that consists of calves between 1985 and 1994;

σ_{p_c} is the standard deviation of p_c^{obs} ;

p_c is the model-estimate of the average fraction of the population that consists of calves between 1985 and 1994;

p_m^{obs} is the observed average fraction of the population that consists of mature animals between 1985 and 1994;

σ_{p_m} is the standard deviation of p_m^{obs} ; and

p_m is the model-estimate of the average fraction of the population that consists of mature animals between 1985 and 1994.

Risk analysis

Forward projections are initialized from the posterior distribution corresponding to the status of the stock at the start of 2000 (walrus) or 2002 (bowheads). Following

Witting and Born (2005), the catch during the first five years of the projection period is set equal to that for during the last year of the assessment (e.g., 80 walrus, with a sex ratio 9:1 M:F), and the population is then projected forward under different levels of constant catch C , for another 5 years (applying the assumed sex-ratio and selectivity pattern of the current hunt). The management objective, ob , used to summarize the results of the decision analysis follows the aboriginal whaling guidelines of the IWC (2000), as interpreted by Witting and Born (2005):

$$ob = \begin{cases} N_{yr+5} > N_{yr} & \text{if } N_{yr} < N_{MSYL} \\ C \leq 0.9MSY & \text{if } N_{yr} \geq N_{MSYL} \end{cases} \quad (2.8)$$

N_{yr} is the population size in 2005 for E-G walrus, and 2007 for B-C-B bowheads. The probability of meeting the objective, given a future catch level and one of the alternative re-sampling schemes, is calculated as the proportion of trajectories at the start of N_{yr+5} that meet the objective. These probabilities are conditioned on reported catch history.

2.3 RESULTS

B-C-B bowhead whales

In general, all three resampling schemes lead to reasonable fits to the data (see, for example, the results for the ‘1978 Fwd’ model in Fig. 2.2). However, there are certain noteworthy differences in the outputs of the models among resampling schemes, especially those from the ‘1978 Fwd’ model. Specifically, resampling f_{\max} , S_a , and a_m consistently leads to higher values for stock productivity, as quantified by λ_{\max} , and hence to better fits to the estimate of abundance for 2001. One consequence of this is that resampling f_{\max} , S_a and a_m leads to the largest estimate of 2002 population size for the ‘1978 Fwd’ model (Fig. 2.2; Table 2.5).

As expected from previous research (e.g. Punt and Butterworth, 1999), the ‘1848 Bkwd’ model is relatively insensitive to modifications to the prior. This is likely the result of conditioning the model on the historical catch record from 1848, which is assumed known without error. There are basically no differences between median estimates of current population size among resampling schemes for this model (although the CV of current population size for the first resampling scheme is appreciably smaller). The posterior median for K is slightly lower when f_{\max} , S_a and a_m are resampled, resulting in a higher probability that the stock is less depleted according to this resampling scheme (Table 2.5; Fig. 2.3).

It is useful to examine the original (explicit) priors, the realized priors and the posteriors for the model parameters to better understand the reasons for the differences among the three resampling schemes for some of the model outputs. Qualitatively, the results for the age-at-maturity are relatively insensitive to the resampling scheme (Fig. 2.4, left panels) with both the realized prior and posterior distributions being unimodal. Closer inspection of the results reveals, however, that resampling f_{\max} , S_a and a_m leads to lower posterior medians than the other schemes. The situation for f_{\max} is similar, with this scheme again leading to the most optimistic posterior (Fig. 2.4, right panels). Likewise, the realized priors and posterior distributions assign less probability to the highest values of adult and juvenile survival when none of the parameters are resampled (scheme ‘None’) (Fig. 2.5).

The realized prior for the maximum rate of increase differs substantially from the explicit prior when no parameters are resampled. Resampling ‘None’ assigns almost no realized prior probability to high (≥ 1.05) values for λ_{\max} , but this is not the case when f_{\max} and S_a , and (particularly) when f_{\max} , S_a and a_m are resampled (Fig. 2.6).

B-C-B bowheads: sensitivity analysis with inflated catches

Fig. 2.7 compares the posterior distributions for the time-trajectory of 1+ population size given the reported catch history (top panel) with that from the analyses in which the catch history is increased fivefold (bottom panel). The estimates of K are obviously very different between catch histories. However, these estimates are effectively insensitive to the choice of resampling scheme for each catch history (Table 2.5). The estimates of the catch quantity Q_1 are only moderately more sensitive to the choice of resampling scheme than those given the reported catch history (Table 2.5).

B-C-B bowheads: sensitivity of risk analysis

There is essentially no difference between resampling schemes in the probability of meeting the management objective (i.e., all predict ~100% success) for future catches as high as 100 whales (Fig. 2.8, left panel). Regardless of the resampling scheme, the resulting prediction is consistent across a wide range of plausible future catch levels. At higher catch levels, resampling ‘None’ leads to more conservative results (Fig. 2.8, left panel).

E-G walrus

The results for E-G walrus illustrate how alternative resampling schemes may potentially impact posterior distributions when the data set is uninformative (Table 2.6). Given the reported catches, the population trajectories for E-G walrus show little sensitivity to alternative resampling schemes (Fig. 2.9, top panel). The median and 90% credibility intervals for N_{2000} are nearly identical for all three schemes. This is perhaps not unexpected because the population is estimated to be at a large fraction of K (Table 2.6). As expected from the bowhead analyses, the results of resampling only a subset of the life history parameters are more similar, and differ from those for resampling ‘None’. Resampling f_{\max} and S_a , and (particularly) f_{\max} , S_a and a_m is again more optimistic in terms of management-related quantities such

as Q_1 . However, unlike the case for the B-C-B bowheads, the impact of the choice of resampling scheme on the management-related quantities can be quite large (e.g. ~50% differences among schemes in the posterior median for Q_1 ; Table 2.6).

The posterior distributions for some of the management-related quantities are nearly identical among the three resampling schemes. For example, the posterior distribution for K is centred around 1000 (slightly higher for ‘None’) and skewed to the right, that for $MSYL$ is very similar to its prior, and the posterior for the catch related quantity Q_1 is skewed to the right and its median differs among resampling schemes (Table 2.6). The sensitivity of the posterior distribution for λ_{\max} to the choice of resampling scheme is consistent with that observed for the B-C-B bowheads, although the size of the effect is much larger for E-G walrus (Table 2.6).

E-G walrus: sensitivity analysis with inflated catches

Fig. 2.9 (bottom panel) shows the posterior distributions for the time-trajectory of 0+ population size from the analyses in which the historical catches are increased fivefold. In contrast to the situation for the ‘low’ catches on which Fig. 2.9 (top panel) was based, the estimates of historical population size are sensitive (i.e., greater than 20% difference in median terms) to the choice of resampling scheme. This sensitivity arises because the catches are now large enough to have reduced the population to well below its carrying capacity; in this situation the stock’s current status does depend on how productive the resource is assessed to be which, in turn, depends on λ_{\max} , and hence the choice of resampling scheme.

E-G walrus: sensitivity of risk analysis

The diverging population trajectories in recent years among resampling schemes for the analyses based on the higher catches (Fig. 2.9, bottom panel) are noteworthy. This pattern is much more pronounced than that for B-C-B bowheads (Fig. 2.7, bottom panel) and has implications for management advice. Therefore, it is not surprising that the results of the risk analysis (e.g., calculating the probability

of achieving a management objective given different levels of catch) are also sensitive to the choice of resampling scheme, even when the analyses are based on the reported (rather than increased) catches. Again, resampling ‘None’ leads to more conservative results (Fig. 2.8, right panel).

2.4 DISCUSSION

These sensitivity analyses suggest that the choice of resampling scheme when implementing the constraint $S_{juv} < S_a$ can impact the results of stock assessments and hence the scientific management advice arising from such assessments. In particular, resampling f_{max} , S_a , and a_m to achieve a near uniform realized prior distribution for λ_{max} consistently leads to more optimistic results given the ranges of the prior distributions considered here (which are representative of many marine mammal populations). The effect can be marked in cases for which the data are uninformative. This is an example of a well known property of Bayesian analyses: when the information content in the likelihood is low, the prior will dominate the resulting posterior. Alternative resampling schemes correspond to different specifications for the realized joint prior, and therefore results of assessments based on limited data are especially sensitive to this issue.

It is noteworthy that the realized prior distribution for λ_{max} when resampling life history parameters assigns higher prior probability to large values for the intrinsic rate of growth. It is also well-known that there are fewer combinations of parameter values for which $S_{juv} < S_a$ when the intrinsic rate of growth is high (Fig. 2.1; Punt and Butterworth 2000). The scheme (‘None’) that does not retain the original value drawn from the prior for λ_{max} in order to find a feasible solution for S_{juv} is therefore assigning greater prior probability to low values for the intrinsic rate of growth. This is because the other two resampling schemes continue to resample the life history parameters when the intrinsic rate of growth is high - to sample that part of

parameter space that satisfies the constraint on S_{juv} - but are not having to do this when the intrinsic rate of growth is low.

In essence, the resampling schemes are a way of re-weighting the marginal priors. The issue is whether one considers the fact that there is less feasible parameter space for larger values of λ_{max} , means that larger values are less likely. Resampling 'None' accepts with equal weight all points of parameter space which respect biological realism. The other extreme is to resample values for all of the life history parameters, which to the maximum extent possible, maintains the prior on λ_{max} . This approach basically ignores the drop in feasible parameter space for larger values of λ_{max} , and appears (for the cases investigated here) to provide the least conservative outcome.

These schemes (except resampling only f_{max} and S_a) have been employed at one time or another during recent assessments of marine mammals: Punt and Butterworth (1999) based their analyses on resampling 'None', and Wade (2002a) resampled f_{max} , S_a , and a_m . The choice among the resampling schemes depends on several factors. The realized prior distributions for the intrinsic rate of growth are nearly uniform when f_{max} , S_a , and a_m or, f_{max} and S_a are resampled (Fig. 2.6). Such distributions are therefore more consistent with the intended prior for λ_{max} . Choosing a resampling scheme that maintains a near uniform prior on λ_{max} is defensible if it is believed that the resampled life-history parameters are essentially nuisance parameters, recognizing that the status of a population depends critically on the value of the parameter that determines productivity, which is λ_{max} in these models. Such an alternative might be appealing, if, for example, there is limited prior information on life history parameters, and instead, there exist a precise series of abundance estimates over a relatively long time-period. In fact, this approach appears to provide a better fit to recent abundance estimates for the B-C-B bowhead stock (Fig. 2.2).

This study was originally motivated by a desire to ensure that, based on the results of recent stock assessments (e.g., Brandon and Wade, 2006) current aboriginal strike limits for the B-C-B bowhead whale are well-founded and sustainable. We emphasize here that this is certainly the case. However, the B-C-B bowhead whale is one of the most well studied populations of marine mammal in the world, and therefore its stock assessments are exceptionally data-rich. This situation is an exception, not a rule. As we have shown, there are other populations of marine mammals for which these issues are an important consideration.

It is not our intention that the results presented here for E-G walrus are directly comparable with those of Witting and Born (2005). The selectivity ogive assumed here for E-G walrus is probably oversimplified (certainly different), and likewise, we explore a catch series that is five times the reported (or estimated) values, purely to illustrate the potential sensitivity of these results. Although the increase in catches is obviously exaggerated, there is undoubtedly considerable uncertainty in this catch history (e.g., Witting and Born (2005) explore a struck and lost rate of up to 25%). It is noteworthy that the estimates of the quantity Q_1 for bowheads are only moderately sensitive to the inflated catch history (Table 2.5). This result is consistent with previous analyses which showed that, given uncertainty or bias in the historic catch record before 1915, the abundance estimates from survey data are the dominant influence on the posterior distributions for quantities related to management (Givens and Thompson, 1996; Givens, 2005).

We have attempted to consistently apply the methods explored here to both case studies. However, it is not practical to use the 'forwards from recent' modeling approach for E-G walrus. There is insufficient data from which to independently specify a prior on recent abundance for this stock and hence fit a population trajectory. If this technique had been used to estimate management quantities for E-G walrus, the results would have been even more sensitive to the specification of priors than was the case for the backwards method. Another difference between the

case-studies is the prior chosen for λ_{\max} . However, both are uniform with an upper bound that is chosen to coincide with the realized upper bound dictated by the constraint on S_{juv} (Fig. 2.1). Although different, the upper bounds do not constrain the results, but do make the numerical integration more efficient. The method investigated here is perhaps the most common approach to constructing a coherent prior for marine mammals in recent years, largely based on assessments performed for the IWC. The consistent methodology allows for comparison between a data-rich and data-poor scenario, highlighting the sensitivity of data-poor scenarios to an easily overlooked aspect of constructing a coherent joint prior while respecting biological realism.

The degree of sensitivity between data-rich and data-poor scenarios is well illustrated by the results of the risk analyses, which are based on the reported catch history. A future catch of 100 bowheads (a catch larger than the current catch) would lead one to predict a consistent ~100% chance of meeting the management objective, regardless of the resampling scheme used (Fig. 2.8, left panel). Whereas, there are large differences in the probability of meeting the management objective for E-G walrus, depending on which resampling scheme is used to construct a coherent joint prior while applying the constraint on juvenile survival (Fig. 2.8, right panel). For example, given an annual catch of 15 animals, resampling ‘None’ indicates that the management objective will be met with only a 60% probability, while resampling f_{\max} and S_a (and a_m) would lead one to predict a greater than 90% probability. If the technique examined here - solving for a free parameter – is to be used to construct a coherent joint prior given a biological constraint, we recommend exploring different resampling schemes during the initial phase of the stock assessment to determine the sensitivity of the results to this choice. Given the results of these analyses, it seems that two coherent joint priors that respect biological realism could be constructed and the assessment run twice, resampling ‘ f_{\max} , S_a and a_m ’ or ‘None.’ This amounts to a sensitivity analysis to alternative joint

priors. These two schemes are likely to bracket intermediately conservative weightings of the marginals. If the results are not consistent between these two schemes, then some consideration should be given to the sensitivity of the assessment results to the priors. We emphasize this recommendation for data poor stock assessments.

Alternative approaches are available to construct a coherent joint prior distribution. For example, a variant of the technique explored here was used by Witting and Born (2005). They imposed a joint prior distribution on (S_{juv}, S_a) , with S_{juv} conditional on S_a such that values of S_{juv} greater than S_a were set equal to S_a . Then, given values of the life history parameters from the prior, they solved for the productivity parameter $MSYR$ (the population growth rate at $MSYL$). Such an approach (setting population growth rate as the free parameter) is appealing if there is good prior information on life history, but not population growth rate. Indeed, walrus populations are an example of this scenario: there exist reliable measurements of life history parameters (e.g. fecundity and age-at-sexual maturity), but accurate surveys of abundance have proven elusive. Solving for $MSYR$ will more closely maintain the explicit priors on life history parameters, for which there is a greater degree of confidence.

Solving for the population growth rate analytically (instead of a life history parameter) does not circumvent biologically impossible solutions (e.g., $\lambda_{\text{max}} < 1.0$). Goodman (1984) clearly demonstrated this fact using Monte Carlo simulation with life history parameters for the spotted dolphin, *Stenella attenuata*. Witting and Born (2005) resampled ‘None’ of the parameters when they arrived at a solution for the maximum population growth rate that was negative. It is just as conceivable to resample only a subset of life history parameters until a feasible solution for population growth rate is obtained. This approach would result in certain realized priors on life history parameters being more consistent with their intended distributions. Choosing to analytically solve for the population growth rate in order

to construct a coherent joint prior, instead of a solving for a life history parameter (e.g., S_{juv}), does not circumvent the need to resample (or ignore) values from the priors when parameter combinations result in solutions for population growth rate that violate biological realism.

Obviously, the population growth rate is an extremely influential parameter with regards to model outputs that are important to management. However, if this parameter is solved for analytically to construct a coherent joint prior, it should be recognized that the resulting implicit prior will be sensitive to the limits placed on the priors for the life history parameters. For instance, given uniform priors on all life history parameters, the resulting implicit prior distribution for population growth rate will be bell-shaped with a mean and variance that shifts according to the upper and lower limits of the uniform priors for life history parameters (c.f. Fig. 1 of Goodman, 1984). So, while very little is often known about certain life history parameters, the range for which this "ignorance" is bounded may be more informative than desired. For example, the resulting implicit prior distribution for the intrinsic population growth rate will differ substantially (all else being equal) between priors for $S_{\text{juv}} \sim U[0.70, 0.90]$ or $\sim U[0.10, 0.90]$. This sensitivity is likely to be unsatisfactory.

If a life history parameter is solved for, then the prior distribution on population growth rate, or at least its upper bounds could be based on a meta-analysis for related species for which there exists trend data from populations recovering from depletion (e.g., Best, 1993). Unfortunately, reliable trend information does not exist for most walrus populations. However, it is worth noting that the constraint on juvenile survival rate, in concert with the ranges of the other life history parameters, is what effectively imposes the upper bound on the realized prior for λ_{max} (Fig. 2.1). Therefore, any uniform prior on λ_{max} will lead to the same results, as long as the upper bound on the explicit prior provides support up to those values for λ_{max} beyond which there is zero realized prior probability. Following the prior

distributions for life history parameters used by Witting and Born (2005), we have chosen a suitable upper limit for E-G walrus $\lambda_{\max} = 1.12$. It is clear from the values for the upper 95th percentiles of the posteriors for this parameter (Table 2.6), that the upper bound on this prior does not constrain the results for E-G walrus.

Setting one parameter to be calculated, given values from the priors of the remaining parameters potentially ignores information. In certain cases, this might be justifiable (e.g., given a lack of knowledge regarding juvenile survival). However, there are other methods for constructing a coherent joint prior while retaining explicit prior distributions on all inputs (and outputs). Poole and Raftery (2000) extend Bayesian synthesis to include logarithmic pooling of priors (French, 1985; Genest and Zidek, 1986). This technique, termed ‘Bayesian melding’, provides a coherent joint prior on model inputs and outputs and is not subject to Borel’s paradox. To our knowledge, this method has only been applied to an age-aggregated surplus production model for the B-C-B stock of bowhead whale (Givens and Robuck, 1999; Poole and Raftery, 2000). It remains to be seen how a biological constraint like $S_{\text{juv}} < S_{\text{a}}$ could be implemented using Bayesian melding with an age-structured PDM. Moreover, it is not obvious how any approach can escape the necessity of resampling (or not) parameter values given that parts of the prior space violate biological realism.

Ideally, whatever method is used to construct a coherent joint prior, an explicit correlation structure between the parameters is involved, given observed and hypothesized relationships among the model parameters (e.g., between fecundity and adult survival). However, completely specifying a joint prior distribution with explicit correlations among the parameters is a complex endeavor for age-structured PDMs. It is interesting to note that implementing the constraint on juvenile survival imposes a correlation structure among the parameters (Fig. 2.10; Punt and Butterworth, 1999), and this is probably an improvement over a naive assumption of independence.

In addition to incorporating correlations among life history parameters, it is desirable that the method used to construct the prior, should change the explicit marginal priors as little as possible. However, it is apparent that resampling (or not) updates the explicit marginal priors, resulting in different realized marginal priors. Further, different resampling schemes change the marginals of the parameters differently. For example, compare the explicit and realized marginals for B-C-B bowhead S_a and λ_{\max} (Fig. 2.5 and Fig. 2.6). Radke *et al.* (2002), in an example from another field of natural resource modeling, achieve a coherent joint prior distribution using Bayesian melding (via SIR), while explicitly incorporating rank correlations among input parameters (see Iman and Conover, 1982; Guan, 2000). The method is analogous to a ‘normal copula’ (e.g., Wang, 2004), which induces a target correlation structure among parameters while retaining the explicit marginal priors. We tested this approach, but it does not appear to be a valid substitution for the constraint on S_{juv} (at least for the cases investigated in these analyses).

It seems certain that the number of assessments using similar Bayesian methods will increase in the future. Our objective is not to advocate a single ‘best’ approach for constructing a coherent joint prior while respecting biological realism. However, as we have shown, Bayesian inference based on assessment scenarios for which data are limited are likely to be quite sensitive to the issues explored here. Likewise, these issues are also relevant to other long-lived marine taxa, for which similar biological assumptions, PDMs and assessment methodologies are appealing, but time-series of abundance and anthropogenic mortality are likely to be limited and imprecise (e.g., some sea-birds, sharks and sea-turtles).

Table 2.1

Estimates, CVs (actually the standard errors of the log abundance estimates, to which these are approximately equal) and the correlation matrix for the indices of abundance for the B-C-B stock of bowhead whales. Source: Zeh and Punt (2005).

Year	Estimate	CV	Correlation matrix									
1978	4765	0.305	1.000									
1980	3885	0.343	0.118	1.000								
1981	4467	0.273	0.056	0.050	1.000							
1982	7395	0.281	0.094	0.084	0.035	1.000						
1983	6573	0.345	0.117	0.104	0.049	0.084	1.000					
1985	5762	0.253	0.070	0.062	0.020	0.078	0.062	1.000				
1986	8917	0.215	0.072	0.064	0.017	0.092	0.064	0.113	1.000			
1987	5298	0.327	0.124	0.110	0.052	0.088	0.110	0.065	0.067	1.000		
1988	6928	0.120	0.028	0.025	0.013	0.017	0.024	0.009	0.007	0.026	1.000	
2001	10545	0.128	0.008	0.007	0.005	0.001	0.007	-0.004	-0.008	0.008	0.003	1.000

Table 2.2

The average proportion of observed calves (p_c^{obs}) and mature (p_m^{obs}) animals, with associated standard errors, over the years 1985 to 1994. Proportions are given based on ignoring the potentially anomalous data set for 1985. Source: IWC (1999) and Koski *et al.* (2006).

p_c^{obs}	σp_c	p_m^{obs}	σp_m
0.0580	0.0062	0.4366	0.0106

Table 2.3

The prior distributions for the B-C-B bowhead whales. Dashes (-) represent prior distributions that are equal to those from the model in the column to the left. 'N/A' indicates a prior that is not required for the model concerned. Fecundity is defined as female calves per mature female. The abbreviations for these distributions are: U ~ Uniform, DU ~ Discrete uniform, and N ~ Normal. Sources are given below.

Parameter		Model Type	
		1848 Bkwd	1978 Fwd
S_a	adult survival	N(0.990, 0.02), truncated at 0.940 and 0.995 ^a	-
f_{\max}	maximum fecundity	U[0.125, 0.200] ^b	-
a_T	age-at-transition to adult survival	DU[1, 9] ^c	-
a_m	age-at-sexual maturity	N(20.0, 3.0) truncated at 13.0 and 26.0 ^d	-
λ_{\max}	intrinsic population growth rate	U[1.005, 1.075] ^e	-
N_{1978}	population size in 1978	N/A	U[3000, 9000] ^f
N_{1993}	population size in 1993	N(7800, 1200) ^g	N/A
K	carrying capacity	N/A	U[8000, 30000] ^h
$MSYL$	$MSYL$ in terms of the 1+ population component	U[0.40, 0.80] ^h	-

a – based on the posterior distribution for adult survival rate obtained by Zeh *et al.* (2002).

b – the prior for the maximum number of calves (of both sexes) per mature female selected by the Scientific Committee of the International Whaling Commission was U[0.25, 0.4] (IWC, 1995). This is the corresponding prior given fecundity has been defined here as female calves per mature female per year.

c – selected by the Scientific Committee of the International Whaling Commission (IWC, 1995) although there is little information on the value of this parameter (Givens *et al.*, 1995).

d – based on a best estimate of 20 years and a lower confidence interval for the age at first parturition (age at sexual maturity + 1 year) of 14 years (IWC, 1995).

e – preliminary analyses indicated there was no posterior probability outside this range, which was confirmed in the final analyses. This range was therefore selected to improve the efficiency of the numerical integration while not affecting the results.

f – selected to encompass a plausible range of values for 1+ population size in 1978.

g – selected by the Scientific Committee of the International Whaling Commission (IWC, 1995) based on the prior distribution assumed for the Bayes empirical Bayes estimate of abundance (Raftery and Zeh, 1991).

h – based on the prior selected by the Scientific Committee of the International Whaling Commission (IWC, 1995).

Table 2.4

The prior distributions for E-G walrus. The abbreviations for these distributions are: U ~ Uniform, DU ~ Discrete uniform and LN ~ Log-normal. Sources are given below.

Parameter		Prior
S_a	adult survival	U[0.900, 0.980] ^a
f_{\max}	Maximum fecundity	U[0.167, 0.250] ^b
a_m	age-at-sexual maturity	DU[5, 9] ^b
λ_{\max}	intrinsic population growth rate	U[1.01, 1.12] ^c
N_{1995}	population size in 1995	LN[ln(1000), 0.35 ²] ^d
$MSYL$	$MSYL$ in terms of the 1+ component	U[0.50, 0.80] ^e

a – prior assumed by Witting and Born (2005), with ranges set wide enough to encompass plausible values as no direct evidence is available for this parameter.

b – Mansfield (1958), Fay (1982) and Born (2001). The range of fecundity values used by Witting and Born (2005) has been divided by two because these values are taken here to relate to the number of female calves per mature female per year.

c – preliminary analyses indicated there was no posterior probability outside this range, which was confirmed in the final analyses. This range was therefore selected to improve the efficiency of the numerical integration while not affecting the results.

d – Born *et al.* (1997). This is the abundance estimate used for 1995 by Witting and Born (2005), where the CV is taken to be approximately equal to the standard error of the logarithm.

e – prior assumed by Witting and Born (2005).

Table 2.5

Posterior medians [5th, 95th percentiles] for five management-related quantities for the B-C-B bowhead population for all models and alternative sampling schemes.

	N_{2002}	K	N_{2002} / K	λ_{\max}	Q_1
1848 Bkwd					
Reported catches					
f_{\max} , S_a , and a_m	9496 [8750, 10180]	10960 [9190, 13950]	0.888 [0.647, 0.985]	1.041 [1.024, 1.059]	228 [149, 296]
f_{\max} and S_a	9571 [8030, 10360]	11670 [9252, 15630]	0.826 [0.459, 0.977]	1.036 [1.014, 1.053]	208 [92, 276]
None	9579 [7974, 10400]	11960 [9562, 16150]	0.809 [0.434, 0.973]	1.034 [1.012, 1.050]	203 [83, 271]
Inflated catches					
f_{\max} , S_a , and a_m	10140 [7957, 11840]	56870 [44750, 66170]	0.180 [0.119, 0.245]	1.058 [1.034, 1.067]	330 [215, 506]
f_{\max} and S_a	9611 [7457, 11410]	58430 [46240, 70070]	0.166 [0.105, 0.226]	1.050 [1.028, 1.067]	290 [177, 441]
None	9364 [7320, 11180]	57550 [46550, 70760]	0.162 [0.101, 0.222]	1.045 [1.026, 1.060]	282 [167, 436]
1978 Fw.					
f_{\max} , S_a , and a_m	10670 [9042, 12410]	20510 [11010, 29120]	0.530 [0.356, 0.925]	1.045 [1.025, 1.063]	295 [160, 439]
f_{\max} and S_a	10210 [7989, 12160]	20890 [9403, 29510]	0.498 [0.302, 0.915]	1.037 [1.015, 1.058]	246 [84, 420]
None	10050 [7949, 11930]	20880 [9253, 29500]	0.487 [0.296, 0.912]	1.034 [1.013, 1.051]	232 [73, 402]

Table 2.6

Posterior medians [5th, 95th percentiles] for five management-related quantities for E-G walrus based on reported and inflated catches.

	N_{2000}	K	N_{2000} / K	λ_{\max}	Q_1
Reported catches					
f_{\max} , S_a , and a_m	1011 [613, 1814]	1067 [661, 1786]	0.98 [0.56, 1.00]	1.059 [1.015, 1.103]	21 [8, 61]
f_{\max} and S_a	1022 [619, 1787]	1082 [681, 1779]	0.98 [0.55, 1.00]	1.053 [1.015, 1.096]	19 [8, 55]
None	1036 [624, 1805]	1176 [751, 1838]	0.91 [0.48, 1.00]	1.033 [1.012, 1.073]	14 [6, 38]
Inflated catches					
f_{\max} , S_a , and a_m	1335 [715, 2361]	3680 [2503, 5906]	0.31 [0.12, 0.72]	1.059 [1.015, 1.103]	50 [10, 122]
f_{\max} and S_a	1312 [701, 2281]	3852 [2627, 5955]	0.29 [0.12, 0.66]	1.053 [1.015, 1.096]	44 [10, 115]
None	1208 [654, 2137]	4575 [3072, 6263]	0.22 [0.10, 0.50]	1.033 [1.012, 1.073]	27 [8, 84]

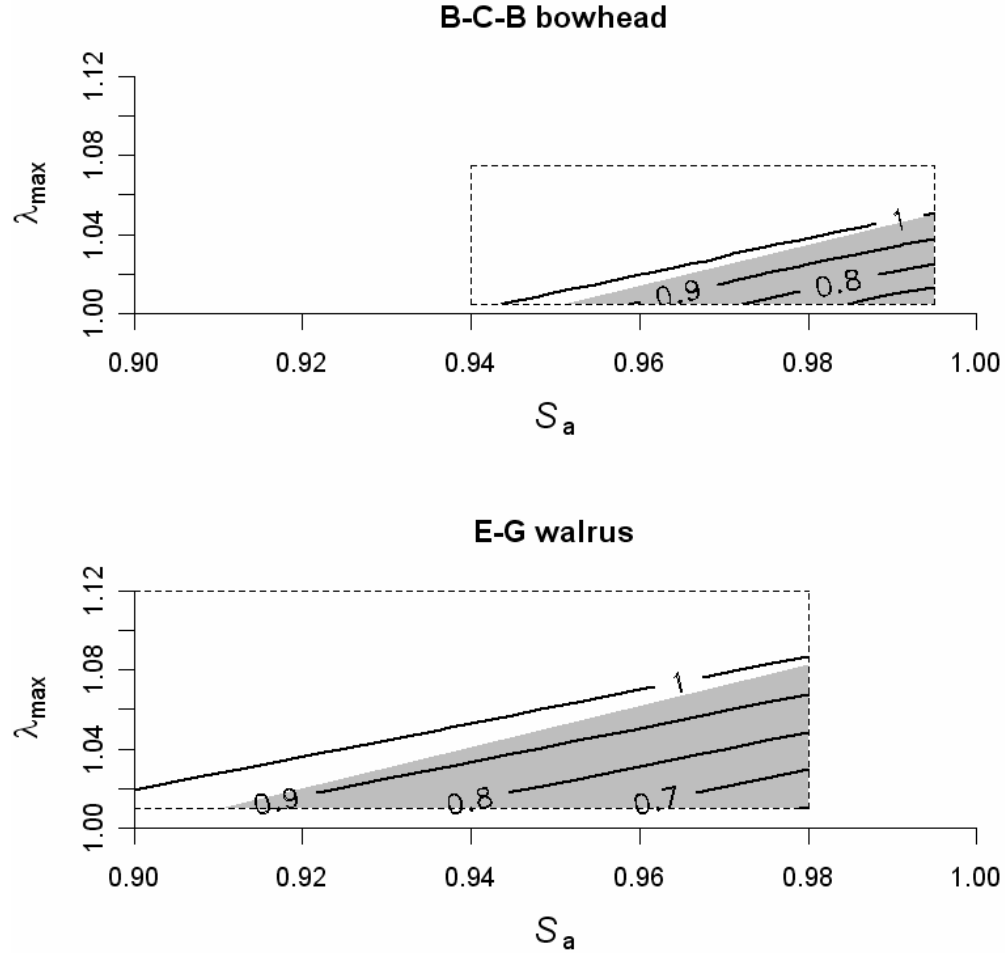


Figure 2.1: Contour plots showing solutions for S_{juv} (from Eqn. 2.4) given a range of values for S_a and λ_{\max} . These diagrams are created by setting the remaining life history parameters constant, equal to the expectation of their explicit prior marginal distribution (e.g., $a_m = 20$ for B-C-B bowheads). The shaded triangular region in the lower right hand corner is the feasible parameter space, subject to the constraint on survival rates. The blank space in the upper left corner represents biologically impossible solutions for S_{juv} which are greater than 1.0. The area between the contour of 1.0 and the shaded region shows the biologically implausible region where $S_{\text{juv}} > S_a$. These plots illustrate the fact that there is more parameter space (given the constraint) consistent with smaller population growth rates.

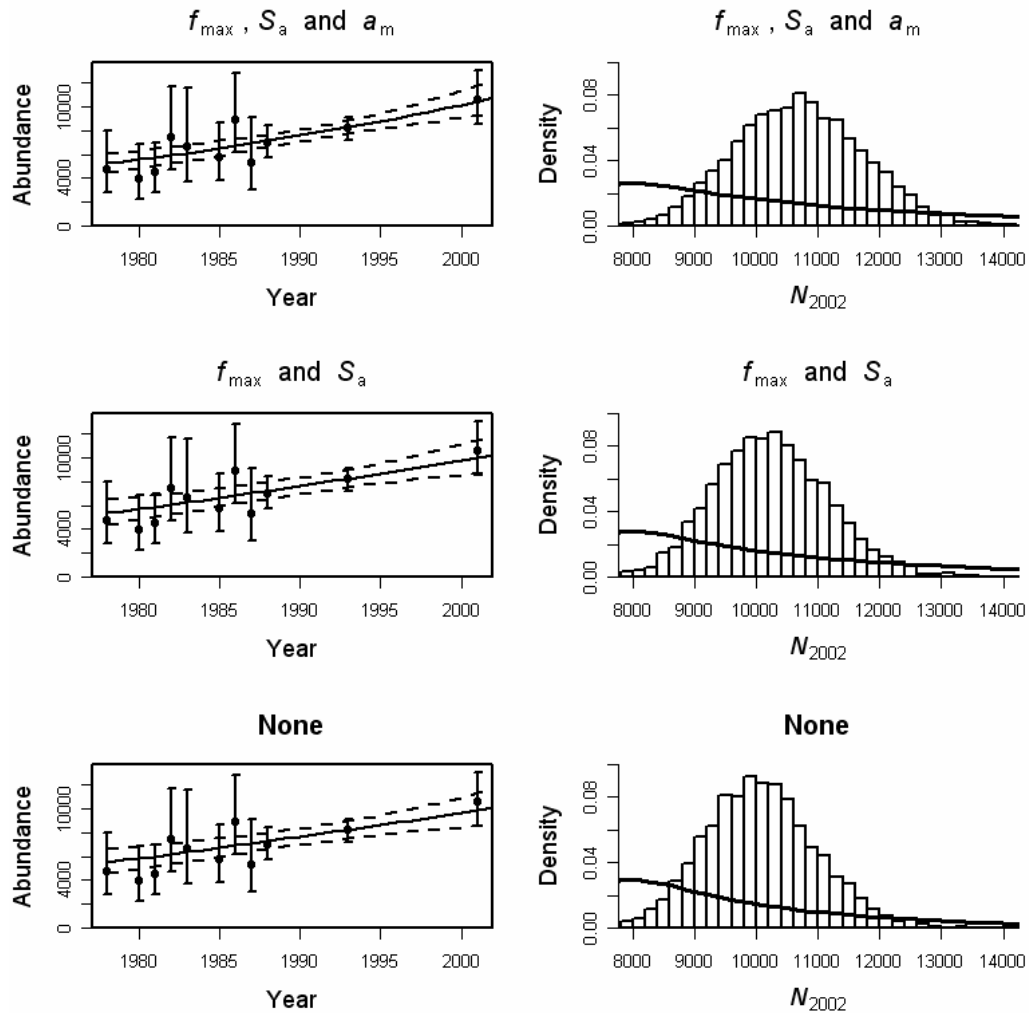


Figure 2.2: Model fits and estimates of recent abundance according to the ‘1978 Fwd’ model for B-C-B bowheads. The left panels show the posterior distributions (medians and 90% credibility intervals) for the time-trajectory of population size. Error bars represent 90% CI’s from survey estimates, and are assumed to be log-normally distributed for all abundance estimates except 1993 (second to last), which is assumed to be normally distributed. The commercial catch during the second half of the 19th century was much higher than the current level, which averages 36 whales during the time period shown. The right panels show the posterior distribution of population size in 2002. The solid lines in the right panels are the realized prior distributions and the bars are the posterior probabilities.

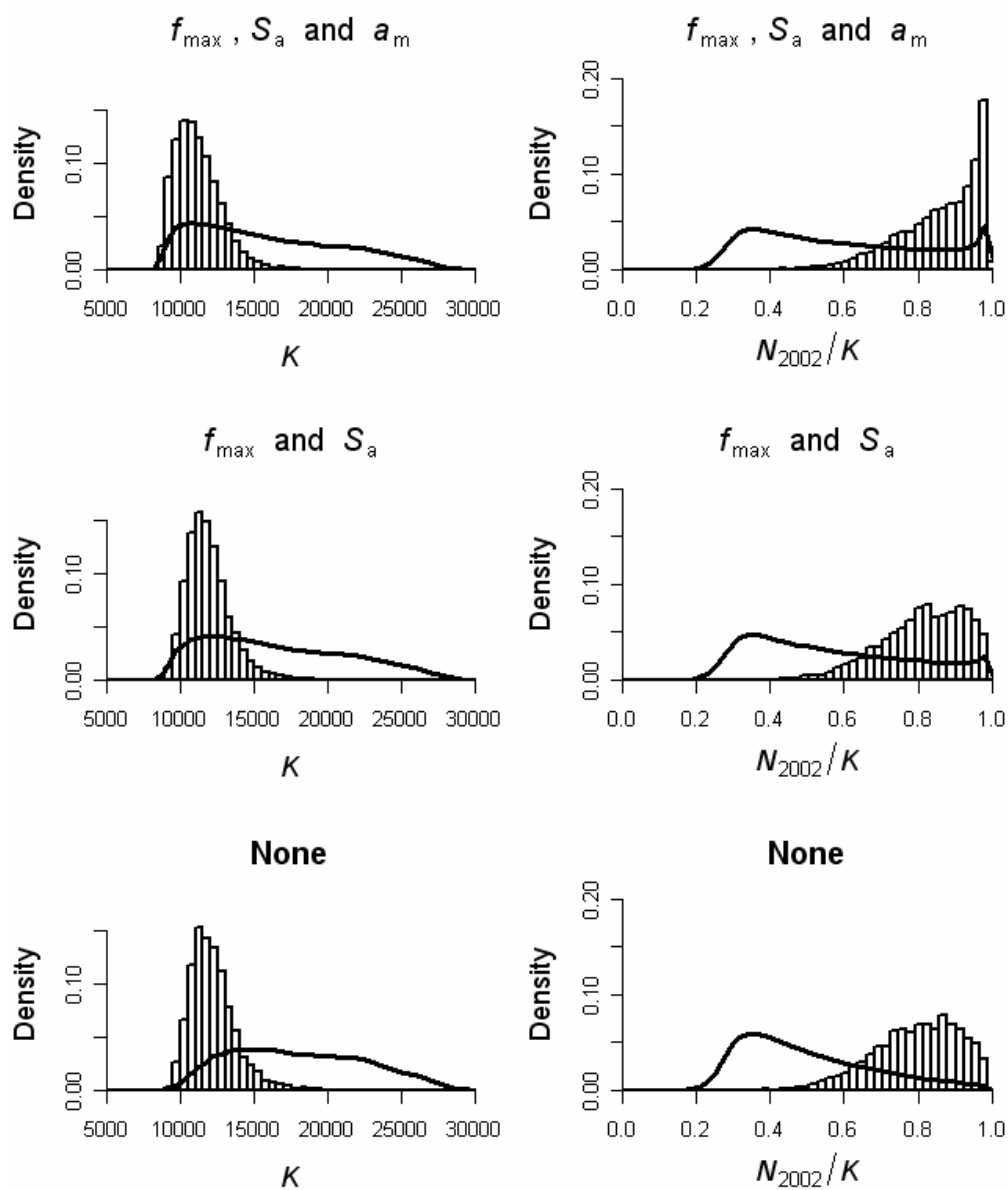


Figure 2.3: Realized prior distributions (solid lines) and posterior distributions (bars) for K and depletion in 2002, according to the '1848 Bkwd' model for the B-C-B bowhead population.

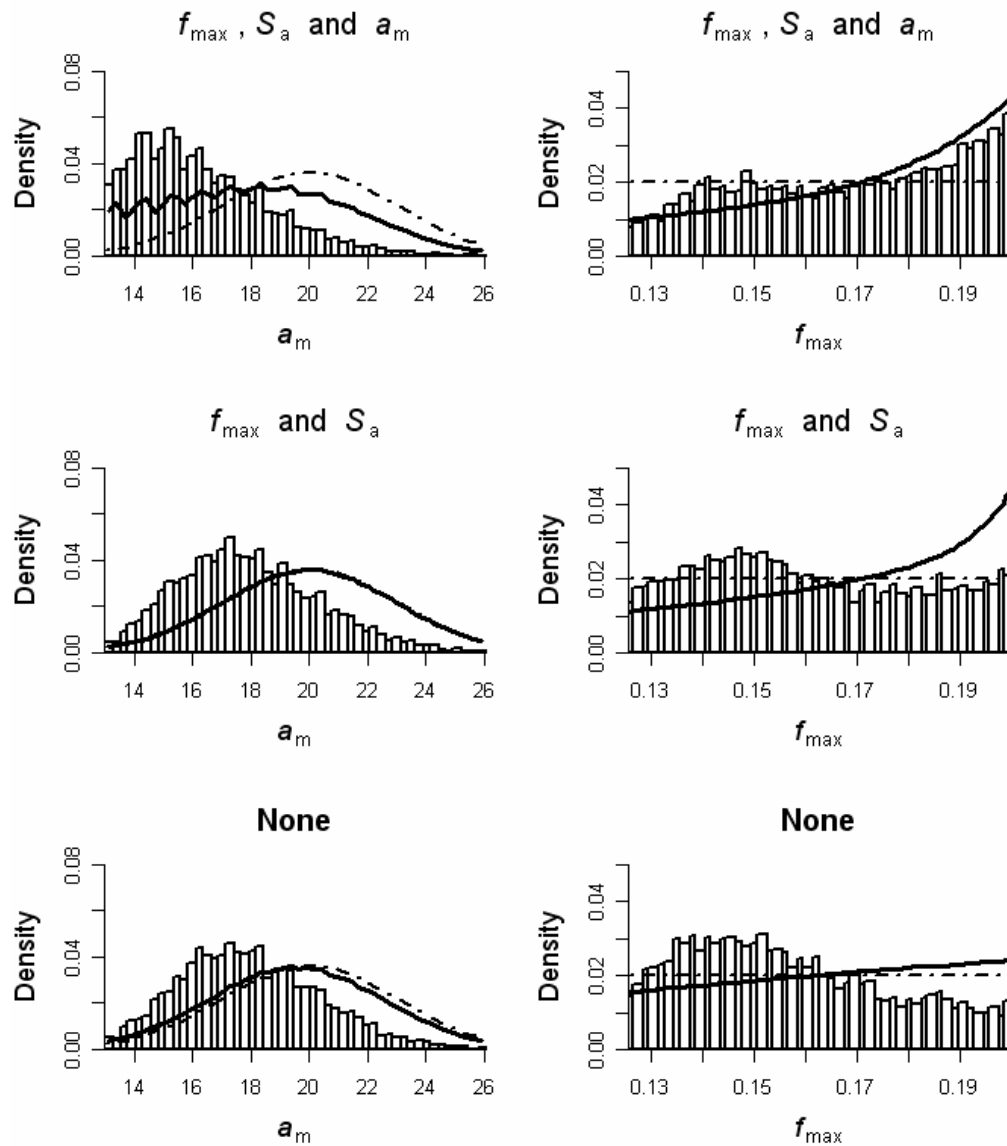


Figure 2.4: Explicit priors (dashed lines), realized priors (solid lines) and posterior distributions (bars) for a_m and f_{\max} , according to the ‘1978 Fwd’ model for the B-C-B bowhead population.

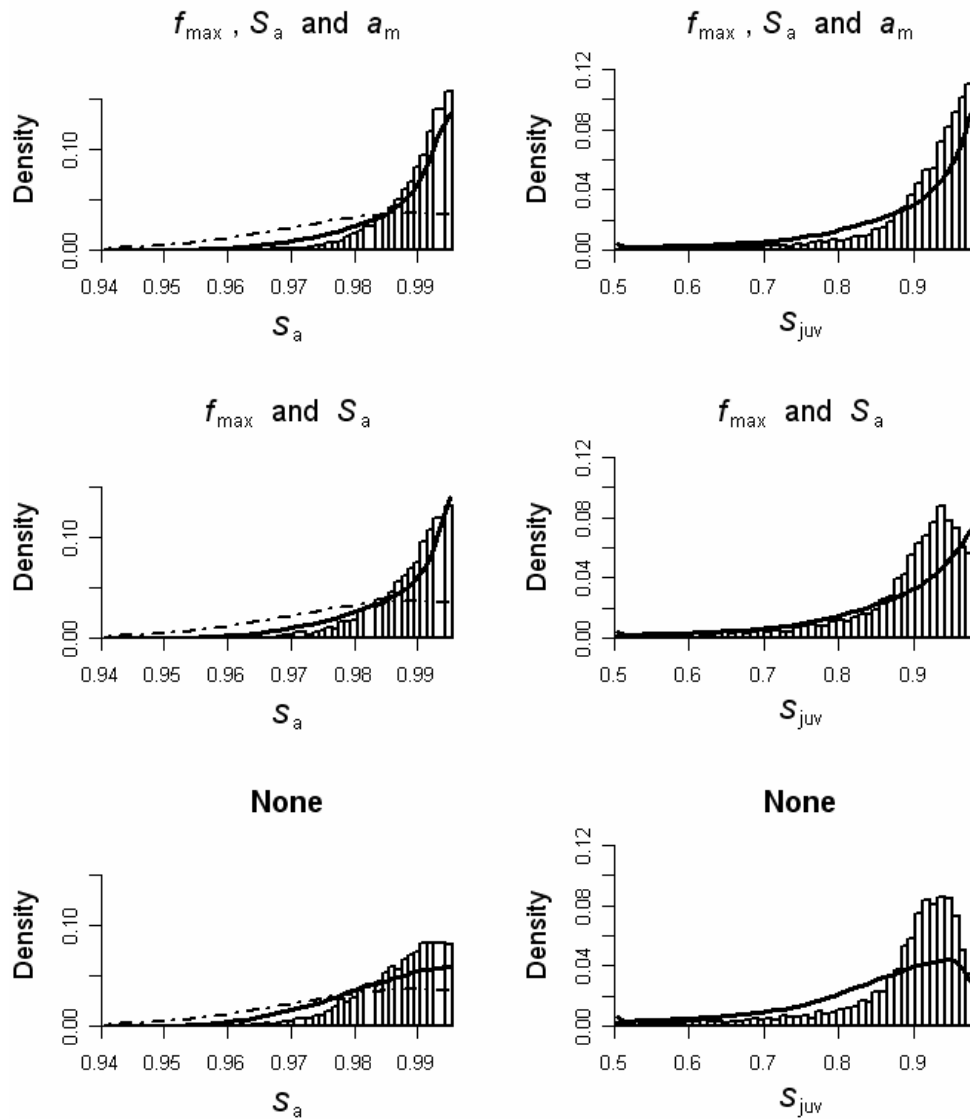


Figure 2.5: Explicit priors (dashed lines – left panels only), realized priors (solid lines) and posterior distributions (bars) for S_a and S_{juv} , according to the ‘1978 Fwd’ model for the B-C-B bowhead population. Note the ranges of the x-axes differ between parameters.

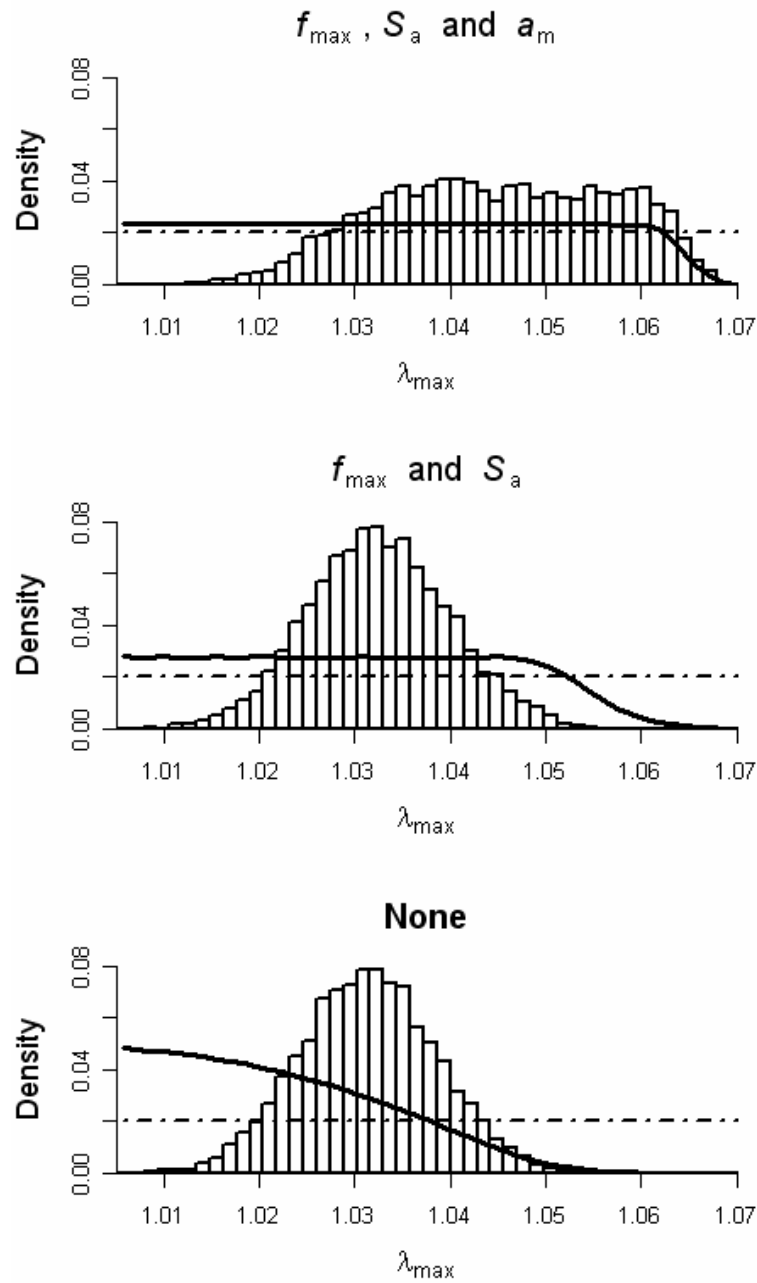


Figure 2.6: Explicit priors (dashed lines), realized priors (solid lines) and posterior distributions (bars) for the intrinsic rate of growth λ_{\max} , according to the ‘1978 Fwd’ model for the B-C-B bowhead population.

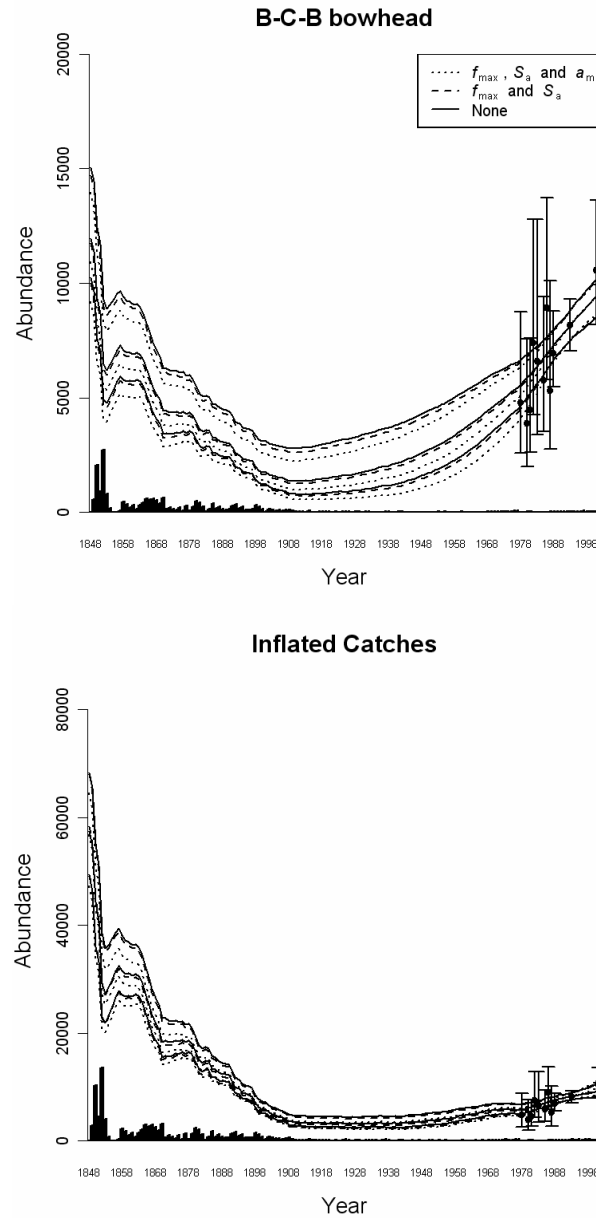


Figure 2.7: Time-trajectories of 1+ abundance (medians and 90% credibility intervals) for the B-C-B bowhead population when the reported catch history is used (top panel) and that when the inflated catch history is used (bottom panel). Note the ranges are different for each y- axis. Catches are plotted along the x-axes. Abundance estimates are shown with error bars representing 90% confidence intervals.

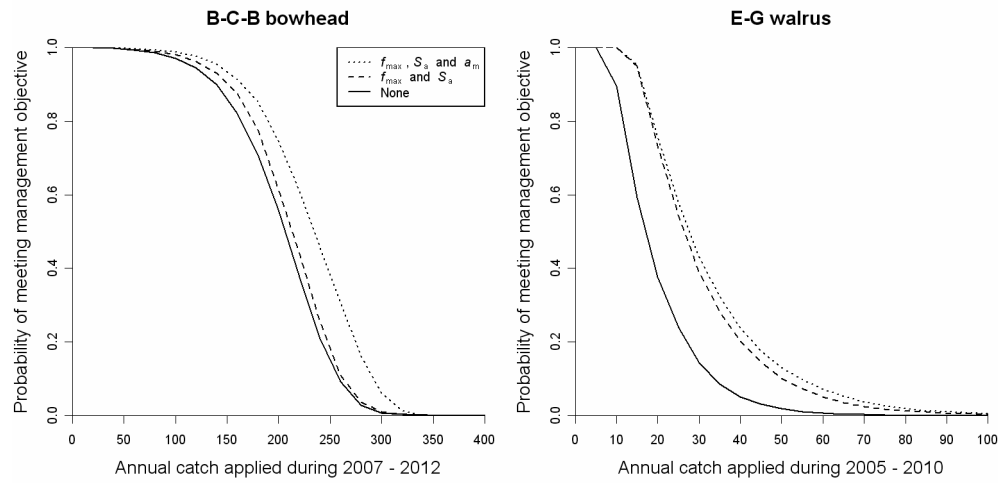


Figure 2.8: The probability of meeting the management objective selected by Witting and Born (2005), for each alternative resampling scheme, as a function of constant future catch for B-C-B bowhead (left panel) and E-G walrus (right panel). Probabilities are shown for each analyses using the reported catch history.

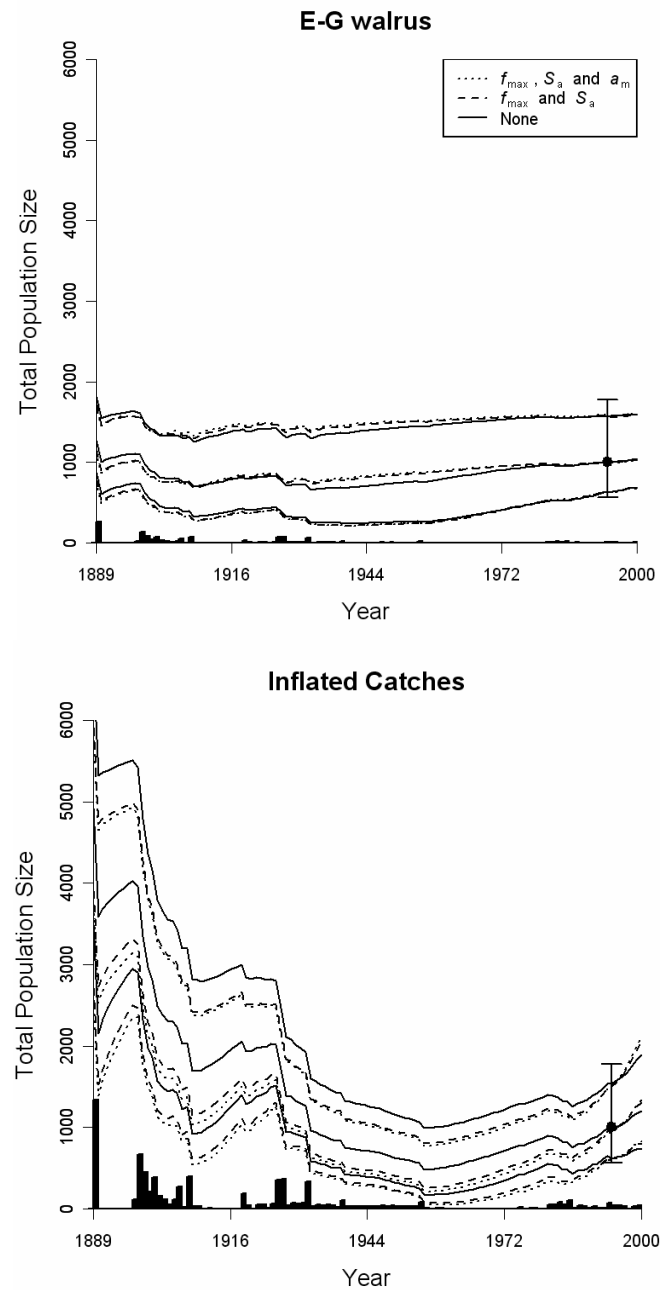


Figure 2.9: Time-trajectories of 0+ population size (medians and 90% credibility intervals) for the E-G walrus population when the reported catch history is used (top panel) and that when the inflated catch history is used (bottom panel). Catches are plotted along the x-axis. The mean of the prior for 1995 abundance (0+) is shown with error bars delimiting the 90th percentiles for the sampling distribution.

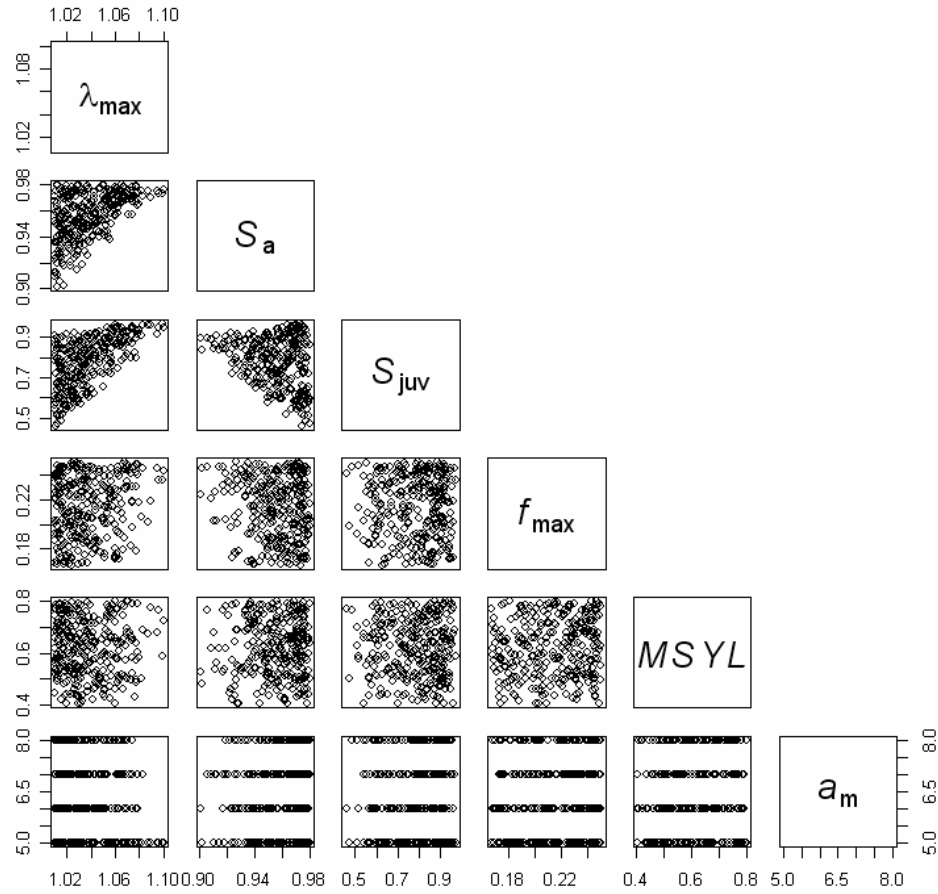


Figure 2.10: Bivariate scatterplots of parameter values from the coherent joint prior achieved by the resampling ‘None’ for E-G walrus.

Chapter 3:

Assessment of the eastern stock of North Pacific gray whales: incorporating calf production, sea-ice and strandings data

ABSTRACT

A stochastic population dynamics modeling framework that integrated a hypothesized relationship between an environmental variable and process error in life history parameters was developed for the eastern North Pacific stock of gray whales. The case study incorporated an index of sea-ice, which has been hypothesized to regulate calf production in this population. The framework also allowed for stochasticity in both birth and survival rates, and was fit to an index of strandings to capture the dynamics observed during the mortality event of 1999 and 2000. Sensitivity tests were performed to evaluate the consequences of various assumptions with respect to the extent of stochasticity and data weightings, and the results of this framework were compared to those based on a deterministic model that was only fit to the abundance data. These alternatives were each able to fit the abundance data well and estimated that the population is very close to the carrying capacity of its environment at present. However, those scenarios which accounted for the effect of the mortality event in 1999 and 2000 led to less optimistic estimates of population status during recent decades with concomitant recovery generally attributable to higher survival rates as opposed to higher birth rates. This study represents the first time that the effects of environmental forcing and the mortality event of 1999 and 2000 have been taken into account during a stock assessment of this population. The framework developed here can be used as an operating model with which to test the Gray Whale *SLA*, given climate forecasts and hypotheses regarding environmental impacts on population dynamics.

3.1 INTRODUCTION

Management of cetacean populations involves determining risk-adverse management strategies that account for natural variability in the environment and impacts of climate change on ecosystems (IWC, 1994; Tynan and DeMaster, 1997). Stock assessment methods that incorporate relationships between environmental factors and population processes offer the potential to improve management in several ways: (1) to increase the precision of parameter estimates for population dynamics models and hence catch or by-catch related quotas (Maunder and Watters, 2003); (2) to provide a tool to evaluate the performance of existing and alternative management strategies given forecasts of future climate (e.g. A'mar *et al.*, 2009); (3) to increase the understanding of factors that might affect the recovery of depleted populations; and (4) to identify priorities for future research and management guidelines (IWC, 1997).

The eastern North Pacific (ENP) stock of gray whales is currently subject to aboriginal hunting, with strike limits based on the Gray Whale Strike Limit Algorithm (SLA) under the Aboriginal Whaling Management Procedure (AWMP) of the International Whaling Commission (IWC 2004, 2005a). The life history of this stock follows a typical baleen whale migration between low and high latitudes (Lockyer, 1984). In general, the majority of animals in the population probably derive most of their annual caloric intake from rich benthic prey communities of the northern Bering and southern Chukchi Seas during the summer feeding season. Measurements of weight and girth support the hypothesis that whales on the northbound leg of the migration have lower fat reserves than their counterparts on the southbound migration, due to reduced feeding during the winter migration (Rice and Wolman, 1971; Perryman and Lynn, 2002). Furthermore, observations from individuals killed off California during the 1960's suggest that pregnant females are the first to migrate northward to the feeding grounds after breeding (Rice and Wolman, 1971). Given these factors and the observed variability in calf

counts during the northbound migration, it has been hypothesized that the extent of sea-ice covering feeding grounds during the early feeding season might act to dictate calf production by affecting feeding opportunities (Perryman *et al.*, 2002).

This chapter provides the first attempt to integrate available estimates of abundance, calf production, strandings and an environmental index that is potentially related to calf production for ENP gray whales in a population dynamics modelling framework. This is one of the most well studied stocks of whales, and therefore provides an ideal candidate to illustrate the results of an assessment method for cetaceans which incorporates an environmental time series. The approach for integrating a relationship between the environmental data and stochastic population dynamics is similar in some respects to that of Maunder and Watters (2003), but differs in that the environmental observations are treated as data and included as a component of the likelihood function (Schirripa *et al.*, 2009). This allows an environmental index with missing years to be incorporated in the analysis. The same approach has been used to assess bigeye tuna (*Thunnus obesus*) in the eastern Pacific (Harley and Maunder, 2004) and is currently being adopted for assessment of sablefish (*Anoplopoma fimbria*) off the US west coast (following, Schirripa and Colbert (2006)).

Unlike previous assessments of this stock, the population dynamics model allows for stochastic birth and survival rates and accounts for three female stages: immature, mature receptive (fertile), and mature with calf. Mature females alternate between calving and receptive stages, with the number of calving females in a given year determined by the stochastic birth rate and the number of receptive females that survived the previous year. This underlying population dynamics model is therefore similar to that of Cooke *et al.* (2007), except that it is age-structured rather than being individual-based. The population dynamics model also attempts to take into account the unusual mortality event that occurred during 1999

and 2000, when anomalously high numbers of individuals were reported dead along the west coast of North America (Gulland *et al.*, 2005).

Results from the application of this approach are compared with those from a deterministic version of the same model, which is not fit to calf production and strandings data, and does not take potential environmental forcing into account. The results of alternative scenarios are also presented for the stochastic model. These scenarios correspond to different assumptions regarding the weights assigned to different data sources or different levels of inherent demographic stochasticity, and allow an evaluation of the sensitivity of the results to key assumptions with respect to these concerns.

3.2 METHODS

Population dynamics model

The analyses were based on a sex- and age-based population dynamics model with an annual time-step. The model included stochastic birth and survival rates, and explicitly considered the transition between receptive and calving stages for mature females (Fig. 3.1). The total number of animals in the population was consequently divided into the number of males $N_{a,t}^{male}$ by age and year, the number of immature females by age and year $N_{a,t}^i$, the number of cows with calves by year $N_{x,t}^c$, and the number of receptive females by year $N_{x,t}^r$.

It was not necessary to explicitly track the age-structure of receptive and calving females because the values for their population dynamic parameters, including harvest rate, were the same for all mature ages. Hence, all females that reached the age-at-maturity were lumped into one of the two mature reproductive stages (denoted as age ‘ x ’). These mature stages are also referred to as the ‘plus-groups’ hereafter, because they included animals which had reached the age at first parturition, plus all animals older than that age. Mature females transitioned

between receptive and calving stages as determined by annual birth rates, with the only exit from these stages through mortality (Fig. 3.1).

Density dependence

Density dependence was assumed to act through the birth rate⁹ according to the Pella-Tomlinson model:

$$b_t = \max \left\{ 0, b_{eq} + (b_{max} - b_{eq}) \left[1 - \left(\frac{N_{1+,t}}{K_{1+}} \right)^z \right] \right\} \quad (3.1)$$

where:

b_{max} is the maximum birth rate (in the limit of zero population size);

K_{1+} is the carrying capacity of the 1+ component of the population (all animals aged 1 yr and older);

b_{eq} is the equilibrium birth rate at carrying capacity;

z is the degree of density-dependent compensation (assumed to equal 2.39, which implies maximum sustainable yield at a population density approximately 60% of K_{1+}), and;

$N_{1+,t}$ is the size of the 1+ component of the population (both sexes combined) in year t .

Stochastic birth and survival rates

Birth rates varied annually about the deterministic value given by Eqn. 3.1. Since this rate must lie between zero and one, its realization in any one year was calculated using a logistic transformation:

⁹ This is really the rate at which receptive females successfully conceive and then survive with calf to make it past central California on the northbound migration. Therefore, this rate will be less than the true birth rate due to early calf mortality, and even more so less than pregnancy rates due to the combined effects of prenatal mortality. Also note that, a constant birth rate of 1.0 corresponds with roughly 50% of mature females having a calf in any given year, due the nature of the population dynamics model.

$$b_t^* = \left[1 + \exp(-(\Phi^{-1}(b_t) \sqrt{2.76 + \sigma_\varepsilon^2} + \varepsilon_t + \varepsilon_{add-1,t})) \right]^{-1} \quad (3.2)$$

where:

- b_t^* is the realized birth rate during year t ;
- Φ^{-1} is the inverse standard normal cumulative distribution function;
- ε_t is the process error deviation during year t , such that $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$;
- σ_ε is a measure of the extent of variability in the process error, and;
- $\varepsilon_{add-1,t}$ allows for additional process error in the birth rate for 1999 and 2000 (in other years, this parameter was set equal to zero).

This formulation of stochastic birth rates ensured that the expected birth rate in a given year was equal to the deterministic value from Eqn. 3.1 (see Appendix A for the derivation, including an explanation of the number 2.76). This transformation leads to a realized standard deviation (taken across years) for the process error deviations that is less than σ_ε (Punt, 2009). Therefore, the realized standard deviation σ_ε was also calculated for comparison.

Natural (i.e. non-fishing) survival rates were also allowed to vary annually with the same process error residuals as birth rates (i.e. the deviations in birth and survival rates were assumed to be perfectly correlated). It was assumed that these rates were independent of sex and perfectly correlated among ages in a given year, such that:

$$S_{a,t}^* = \left[1 + \exp(-(\Phi^{-1}(S_a) \sqrt{2.76 + \sigma_\varepsilon^2} + \varepsilon_t + \varepsilon_{add-2,t})) \right]^{-1} \quad (3.3)$$

where:

- $S_{a,t}^*$ is the realized age-specific survival rate during year t ;

S_a is the deterministic survival rate from age a to $a+1$, and;
 $\varepsilon_{add-2,t}$ is a parameter which allows for additional process error in survival rates in 1999 and 2000 (in other years, this parameter was set equal to zero).

Note, only two natural survival rates were modeled in these analyses: (i) calf survival (S_0) to age 1; and (ii) the survival rate of animals aged 1 year and older (S_{1+}). That is, survival rates were assumed equal for all animals aged 1 and older (i.e., $S_{a>0} = S_{1+}$). Additional details are provided below, under ‘Parameterization and scenarios’.

Preliminary analyses indicated that the distributional assumption for the process error deviations did not allow the model to fit the relatively extreme observations of strandings during 1999 and 2000. Thus, ε_{add-1} and ε_{add-2} were introduced into Eqns. 3.2 and 3.3 for those years to try and capture the population dynamics during the mortality event of 1999 and 2000. Two sets of scenarios were run: 1) $\varepsilon_{add-1} = \varepsilon_{add-2}$ (the base-case), and; 2) ε_{add-1} and ε_{add-2} estimated individually (Table 3.1).

Female dynamics

The number of immature females by age depended on the number of births, an assumed 50:50 sex ratio at birth, maturation, and mortality from natural causes and hunting. Maturity was assumed to be knife-edged at age 6 (i.e. all females reached the age at first estrous at age 6). The gestation period was assumed to be one year, so the age at first possible parturition was 7 yr, which is equivalent to the median of the prior distribution for this life history parameter adopted in previous assessments of the ENP gray whales (IWC, 1993; Wade, 2002a). The plus-group age was set

equal to the assumed age at first parturition (i.e., $x=7$ years or older, i.e., 7+), so there was no need to implement a plus-group for the immature stage:

$$N_{a,t+1}^i = \begin{cases} 0.5N_{t+1}^c & \text{if } a = 0 \\ N_{a-1,t}^i S_{a-1,t}^* \left(1 - E_t^{fem} V_{a-1}\right) & \text{if } 1 \leq a \leq 6 \end{cases} \quad (3.4)$$

where

E_t^{fem} is the exploitation rate during year t on females:

$$E_t^{fem} = C_t^{fem} / \left(N_{x,t}^r + N_{x,t}^c + \sum_a V_a N_{a,t}^i \right) \quad (3.5)$$

V_a is the selectivity on animals of age a , assumed to be constant with regard to sex and time, and uniform on ages 5+, following the approach of previous IWC assessments (IWC, 1993):

$$V_a = \begin{cases} 0 & \text{if } a < 5 \\ 1 & \text{if } a \geq 5 \end{cases} \quad (3.6)$$

C_t^{fem} is the total catch of females during year t .

This formulation assumed that selectivity was the same for all animals of a given age, and was independent of sex, time, and reproductive condition. Hence, the assumed selectivity pattern allowed for cows, but not calves to be killed in the hunt.

The number of receptive females and cows with calves was (given full selectivity on these plus-groups) is given by:

$$N_{x,t+1}^r = S_{1+,t}^* \left[\left(1 - b_t^*\right) \left(N_{6,t}^i + N_{x,t}^r \right) + N_{x,t}^c \right] \left(1 - E_t^{fem}\right) \quad (3.7)$$

$$N_{x,t+1}^c = S_{1+,t}^* b_t^* \left(N_{6,t}^i + N_{x,t}^r \right) \left(1 - E_t^{fem} \right) \quad (3.8)$$

Male dynamics

Males were modeled using an age-structured model that ignored maturity because the number of males was assumed not to be a limiting factor for female reproductive success. However, plus-group dynamics ($x=7+$) were modeled for males because this avoided the assumption that all animals died after a certain age (given full selectivity on the plus-group):

$$N_{a,t+1}^{male} = \begin{cases} 0.5N_{t+1}^c & \text{if } a = 0 \\ N_{a-1,t}^{male} S_{a-1,t}^* \left(1 - E_t^{male} V_{a-1} \right) & \text{if } 1 \leq a < x \\ S_{1+,t}^* \left(1 - E_t^{male} \right) \left(N_{6,t}^{male} + N_{x,t}^{male} \right) & \text{if } a = x \end{cases} \quad (3.9)$$

where:

E_t^{male} is the exploitation rate during year t on males:

$$E_t^{male} = C_t^{male} / \sum_a V_a N_{a,t}^{male} \quad (3.10)$$

C_t^{male} is the total catch of males during year t .

Initial conditions

Population trajectories were initiated in 1930, under the assumption of a stable-age-distribution given some level of hunting mortality in 1930. A numbers-per-female-calf approach was taken to solve for the numbers-at-age in 1930 given values for the life-history parameters of the model, the relative size of the 1+ component in 1930, and the hunting mortality rate in 1930, E_{init} . The number of females per calf is given by (given full selectivity on the plus-group):

$$NPR_a^{(E)} = \begin{cases} 0.50 & \text{if } a = 0 \\ NPR_{a-1}^{(E)} S_{a-1} (1 - E_{\text{init}} V_{a-1}) & \text{if } 1 \leq a < x \\ NPR_{x-1}^{(E)} S_{1+} (1 - E_{\text{init}}) / (1 - S_{1+} (1 - E_{\text{init}})) & \text{if } a = x \end{cases} \quad (3.11)$$

The birth rate at unexploited equilibrium b_{eq} is the inverse of the number of receptive females per-calf which can give birth. Since the maturity ogive was assumed to be knife-edged and the age at first parturition was assumed equal to the age at which individuals entered the plus group, the number of mature females-per-calf was $NPR_x^{(E=0)}$. Given this, b_{eq} is:

$$b_{eq} = \left(NPR_x^{(E=0)} - 1 \right)^{-1} \quad (3.12)$$

The numbers-per-recruit approach of Punt (1999a) was modified to take account of hunting mortality in 1930. This involved calculating b_{1930} using Eqn. 3.1 given b_{eq} , b_{max} and the relative size of the 1+ component in 1930, and using Newton's method (Press *et al.*, 1992) to solve for the value of E_{init} such that:

$$1 = b_{1930} \left(NPR_x^{(E=E_{\text{init}})} - 1 \right) \quad (3.13)$$

The age- and sex-structure at the start of the 1930 was then calculated by scaling the numbers-per-calf by the number of calves corresponding to K_{1+} . The numbers-at-age of each sex in 1930 was then the total numbers-at-age divided by two.

Data and likelihood function

Four sources of data were considered when fitting the full model: (1) estimates of population size during 1967-2006 (starting year of survey) from the southbound migration at Granite Canyon, California (Rugh *et al.*, 2005, 2008); (2) estimates of

calf production during 1994-2008¹⁰ from the northbound migration at Point Piedras Blancas, California (Perryman *et al.*, 2002; Perryman, *unpublished data*), (3) the number of stranded animals on the coasts of California, Oregon and Washington state, for which a combined annual count was available during 1975-2006 (Brownell Jr. *et al.*, 2007)¹¹, and; (4) estimated sea-ice area covering the Bering Sea, averaged over March and April during 1953-2006 as calculated by the Hadley Center for their sea-ice and sea surface temperature data set version 1 ('HadSST') (Fig. 3.2, left panel; Rayner *et al.*, 2003).

The HadSST ice index was used here because it represents a good compromise between a shorter high-resolution and a longer less-precise environmental index. It is a compilation of several sources of data including the 'Walsh' charts of sea-ice extent prior to 1978 (Walsh, 1978), and satellite observations for recent decades. Therefore, this index provides a relatively long time series, calibrated by recent satellite observations. It is also worth noting that one of the major goals of these analyses is to use those results to test the Gray Whale *SLA* given predictions of future sea-ice. Since the primary purpose of the HadSST index is to form the basis for forcing atmospheric circulation models during simulations of future climate, and because we plan on using the results from such simulations of future climate when testing the Gray Whale *SLA*, the HadSST is the sea-ice index most consistent with the objectives of this research.

Catches by sex are available from 1930-2006 (Fig. 3.2, right panel). Selectivity-at-age resulting in the observed catches was treated as known. Hence no attempt was made to fit the catch data. Instead, catches were simply subtracted from the population each year according to the assumed selectivity ogive. The catches during 2007 and 2008 were assumed equal to those in 2006.

¹⁰ The two early estimates of calf production during 1980-1981 (Poole, 1984) were not used in these analyses, as they are not currently used when testing the Gray Whale *SLA*.

¹¹ Data on strandings are collected in other locations (e.g. Mexico and Alaska), but the stranding network effort in California, Oregon and Washington has been more consistent through the years

The total negative of the logarithm of the likelihood function is the sum of the contributions for each data source. In addition, penalties were added to the likelihood function to impose a normal prior with standard deviation σ_ε on the process error deviates and to ensure that trajectories resulting in extinction were assigned zero likelihood. Previous studies using similar, but not identical methods of including process error in the population dynamics have assigned values for σ_ε using an approach which relies on the convergence of the root-mean-squared-error between the logarithms of expected vs. observed recruitment (Methot, 2000). However, that approach is not suitable given the transformation applied here (Equations 3.2 and 3.3). Instead, a default value for σ_ε of 0.50 was used following preliminary analyses which suggested that this value was consistent with observed inter-annual variability in the data. Analyses were also conducted in which $\sigma_\varepsilon = 0.30$ and 0.70 to assess the sensitivity of the results to the value assumed for σ_ε .

Abundance estimates

The abundance estimates are based on survey seasons which span two calendar years. They are referred to here by the year during which the survey started (e.g., the 1967-68 abundance estimate is referred to as “1967”). In this way, the abundance of the population was considered to be surveyed after births and deaths in a given year. An additional variance term $CV_{\text{add-1}}$ was incorporated into the likelihood component for the abundance estimates following Wade (2002a) and Butterworth *et al.* (1993). The residuals of the fit to the abundance estimates were assumed to be independent between years and log-normally distributed. Moreover, it was assumed that the surveys provided estimates of the number of animals aged 1

and older. These assumptions led to the following component of the negative log-likelihood function¹²:

$$L_1 = \sum_t 0.5 \left(\ln(\sigma_t^2 + CV_{\text{add-1}}^2) + \frac{1}{\sigma_t^2 + CV_{\text{add-1}}^2} \left(\ln N_{1+,t}^{\text{obs}} - \ln N_{1+,t} \right)^2 \right) \quad (3.14)$$

where:

$N_{1+,t}^{\text{obs}}$ is the survey estimate of 1+ abundance for year t ;

$N_{1+,t}$ is the model estimate of 1+ abundance for year t ;

$CV_{\text{add-1}}$ is the extent of additional error about the abundance estimates, and;

σ_t is the standard deviation of the logarithm of $N_{1+,t}^{\text{obs}}$ (approximated by the CV of the untransformed abundance estimate).

Calf estimates

The residuals about the model fit to the calf estimates were also assumed to be independent and identically log-normally distributed. Following previous approaches which fit the calf estimates using a deterministic population dynamics model (Wade, 1997; Wade and Perryman, 2002), the reported observation error about the calf estimates was assumed to be subject to some additional observation error as was the case for the abundance estimates. This approach led to the following component of the negative log-likelihood function:

$$L_2 = \sum_t 0.5 \left[\ln(\sigma_t^2 + CV_{\text{add-2}}^2) + \frac{1}{\sigma_t^2 + CV_{\text{add-2}}^2} \left(\ln N_{0,t}^{\text{obs}} - \ln N_{0,t} \right)^2 \right] \quad (3.15)$$

where:

$N_{0,t}^{\text{obs}}$ is the survey estimate of calf production in year t ;

¹² The likelihood components were calculated ignoring constants independent of the parameters of the model.

- $N_{0,t}$ is the model estimate of calf production in year t ;
- $CV_{\text{add-2}}$ is the extent of additional error about the calf estimates, and;
- σ_t is the standard deviation of the logarithm of $N_{0,t}^{\text{obs}}$ (approximated by the CV of the untransformed calf production estimate).

Stranding counts

The residuals about the model fit to the indices of number of stranded animals were assumed to be independent and identically log-normally distributed, leading to the following component of the negative log-likelihood function:

$$L_3 = \sum_t 0.5 \left[\ln(\sigma_M^2) + \frac{1}{\sigma_M^2} \left(\ln M_t^{\text{obs}} - \ln(\hat{q} \hat{M}_t) \right)^2 \right] \quad (3.16)$$

where:

- \hat{q} is the constant of proportionality between the indices of stranded animals and \hat{M}_t ;
- M_t^{obs} is the observed number of stranded animals (based on data for California, Oregon and Washington), and;
- \hat{M}_t is the model-estimate of the number of animals (for all stages and both sexes) dying due to natural causes:

$$\hat{M}_t = (1 - S_{1+,t}^*) [N_t^r + N_t^c] + \sum_a (1 - S_{a,t}^*) [N_{a,t}^i + N_{a,t}^{\text{male}}] \quad (3.17)$$

An empirical estimate for the observation error of the stranding counts does not exist (Brownell Jr. *et al.*, 2007). Therefore, reasonable alternative values were chosen ($\sigma_M = 0.10$ or 0.20) to assess the sensitivity of the results to the value assumed for this parameter. A value for σ_M of 0.20 implies that the lower 95% limits for the stranding estimates for 1999/2000 do not overlap with the upper 95%

limits for the strandings estimates for any other years, and it is therefore an upper limit for this parameter which would be consistent with those years representing an unusual mortality event.

The value for \hat{q} was set to its maximum likelihood estimate. This is equivalent to integrating over the prior for this parameter when its prior distribution is uniform in log-space (Walters and Ludwig, 1994). It was reasonable to assume that \hat{q} was less than 1.0, because counts of stranding animals were only made along a portion of the migratory route, and further it seems unlikely that all animals that die will wash ashore or that all of those that do will be counted. In addition, an underlying assumption of this method is that \hat{q} was constant through time. This is unlikely to be strictly true. However, given that gray whales migrate (and die) close to the coast, observation effort has been relatively constant through time for the stranding index considered here and that the mortality event of 1999 and 2000 is believed to have been caused by a substantial decrease in survival (as opposed to a higher fraction of carcasses washing ashore due to a change in wind, ocean currents or the like), minor violations of this assumption were unlikely to be consequential to the results.

Environmental impact on demographic rates

In addition to being subjected to process error, the deviations of birth and survival rates about the deterministic relationship each year were also allowed to be related to an environmental index I_t (in this case, the amount of sea-ice covering the Bering Sea, averaged over March and April). It was assumed that I_t was measured subject to observation error (or there was some error in the relationship between the process error deviations, ε_t , and the environmental index). Consequently, I_t was treated as a state variable, similar to the model prediction of population size. The measurements of the environmental index were therefore treated as data and were

consequently included as a component of the likelihood function when the model was fit. The expected environmental index in a given year was assumed to be related to process error residuals for that year, such that the observed index was normally distributed about its expectation:

$$I_t^{obs} = \beta \varepsilon_t + \gamma_t \quad (3.18)$$

where:

- I_t^{obs} is the observed value of the environmental index in year t ;
- β is a scaling parameter for the influence of the environment on the process error residuals;
- γ_t the difference between the observed and model-predicted amount of sea ice in year t , such that $\gamma_t \sim N(0; \sigma_\gamma^2)$, and;
- σ_γ is the standard deviation of the residual error for the environmental index:

$$\sigma_\gamma = |\beta| \sigma_\gamma^* \quad (3.19)$$

This formulation takes a fixed input value for σ_γ^* (Table 3.1) and scales the expected standard deviation of the fits to the environmental index by the estimated absolute value for β . It was found through preliminary analyses that, simply fixing σ_γ to a given fixed input value (ignoring Eqn. 3.19) led to estimates of the process error deviations ε_t which became increasingly small with smaller assumed values of σ_γ . Hence, Eqn. 3.19 leads to the desired effect of the process error deviations being more correlated with the environmental index at smaller values of σ_γ .

Perryman *et al.* (2002) investigated two different time lags (corresponding to the potential effect of sea-ice on ovulation or pregnancy rates) and concluded that a

relationship between sea-ice during a given year and reproductive success would most likely result from an effect on the pregnancy rates in that year. Therefore, the timing of the potential effect of sea-ice variability in a given year was assumed to be related to deviations from expected birth rates in that year as opposed to the previous year (i.e. a potential effect on ovulation rates). Given the estimation framework here, negative values of β correspond with larger values of the sea-ice index having detrimental affects on birth and survival rates (negative process error deviations).

Given the above assumptions, the contribution of the environmental index to the likelihood function was:

$$L_4 = \sum_t \left[\ln(\sigma_I) + \frac{1}{2\sigma_I^2} (I_t^{obs} - I_t)^2 \right] \quad (3.20)$$

σ_I^* was assumed to be 0.30 for the base case scenario, because preliminary analyses indicated that this value provided a conservative weight for the environmental index during the model fitting (i.e., it led to a reasonable balance between not over-fitting the environmental index, while still allowing for a relatively strong signal in the process error deviations). Analyses were also conducted with $\sigma_I^* = 0.10$ and 1.00 to investigate the sensitivity of the results to alternative values. Likewise, two scenarios were considered in which the model was fit only to data for sea-ice pertaining to those years for which it would have had an effect on recent calf production (1993 – 2008). In these scenarios an alternative index of sea-ice was also fit, based on an updated version of the index used by Perryman *et al.* (2002) (Fig. 3.2, left panel; Perryman, *unpublished data*). This was done to assess the impact of the length of the time-series of environmental data on the results, as well as that given for alternative index of sea-ice.

Parameterization and scenarios

The estimable parameters of the population dynamics model are listed in Table 3.2. Rather than treating all of the survival rates by age as estimable parameters, two survival rates were considered: (i) calf survival S_0 and, (ii) the survival rate for animals aged 1 and older S_{1+} . Moreover, calf survival was not treated as an estimable parameter. Instead, the difference, Δ , between adult and calf survival was estimated. This also allowed the constraint that adult survival cannot be less than calf survival to be enforced. All but one of the scenarios in which the calf data were used to fit the model involved setting the level of additional observation error equal to that for the abundance data (i.e. only one CV_{add} was estimated, such that $CV_{add-2} = CV_{add-1}$).

Table 3.1 outlines the full set of scenarios. The two base case scenarios were: (i) the stochastic model described above (“Full” in Table 3.1), and; (ii) a deterministic version fit only to abundance data following the approach of previous assessments (e.g., Wade, 2002a) (“Deterministic” in Table 3.1). Several alternative scenarios were considered for the full stochastic model to investigate the affects of certain assumptions and data sources on the results. These alternative scenarios involved estimating the maximum likelihood estimates for the parameters (MLE), while the two base cases involved parameter estimation using maximum likelihood as well as a Bayesian framework. The latter facilitated comparison with previous assessments and forms a basis for evaluating the performance of the Gray Whale *SLA*. A parallel set of scenarios involved estimating separate values for ε_{add-1} and ε_{add-2} . This was done to assess the ability of the model to fit the mortality event when the process error deviations were not assumed to be equal for birth and survival rates during the mortality event.

Parameter estimation

The models were developed using AD Model Builder (ADMB, Otter Research, <http://otter-rsch.com/admodel.htm>). ADMB uses automatic differentiation (Griewank and Corliss, 1991) to efficiently estimate the variance-covariance matrix of model parameters with respect to the likelihood function. Additionally, it allows for Bayesian estimation by sampling from the posterior distribution using Markov Chain Monte Carlo (MCMC), as implemented by the Metropolis-Hastings algorithm (Hastings, 1970; Gelman *et al.*, 2004). The proposal (or “jump”) function used by ADMB for the MCMC algorithm is multivariate normal with a variance-covariance matrix based on that estimated for the model parameters.

The Bayesian Output Analysis Program (BOA) for MCMC was used to diagnose the convergence of the MCMC algorithm (Smith, 2007). The Heidelberger and Welch (1983) stationarity and half-width tests, and the Geweke (1992) and the Raftery and Lewis (1992) convergence diagnostics were inspected for signs of non-convergence and used as guidelines for determining an appropriate burn-in and thinning interval for the chain.

3.2 RESULTS

The MCMC algorithm was run for 50 million iterations, saving every 25,000th sample after a 20% burn-in, leading to a final sample size of 1,601 draws from the posterior. This process resulted in diagnostics for the chain that gave no sign of not having converged, as indicated by Figure 3.3.

The model was able to fit the abundance and calf data reasonably well for all scenarios (see Fig. 3.4 for three examples for calf data; Fig. 3.5 upper panels for abundance data for all scenarios). The results were consistent with this stock being at or near carrying capacity, although estimates of carrying capacity differed among scenarios (Fig. 3.6, left panels; Tables 3.3 and 3.4). The scenarios that did not take the strandings data into account (the ‘Deterministic’ and ‘No Strandings Data’

scenarios), or did not place much weight on the strandings data (the $\sigma_M = 0.20$ scenario) estimated carrying capacity to be in the low 20,000s, and that the population size has been constant at this level since the late-1980s or early 1990s (Fig. 3.4, upper right panel; Fig. 3.5, lower panels). The inability of certain scenarios to fit the 1999-2000 mortality event is indicated by values of ε_{add-2} that are closer to zero in Tables 3.3 and 3.4. The estimates of maximum birth rate and survival rates were similar among the scenarios that were unable to fit the 1999-2000 mortality event, with higher maximum birth rates and somewhat lower survival rates than for the remaining scenarios (Tables 3.3 and 3.4). The “Full” scenario estimated life history parameters more precisely than the “Deterministic” scenario, and also estimated lower maximum birth rates and higher survival rates (Fig. 3.7).

The scenarios which estimated both CV_{add-1} (abundance) and CV_{add-2} (calf) resulted in estimates for CV_{add-1} which were generally equal to those for the other scenarios. However, CV_{add-2} was estimated to be equal to zero. The results from estimating this additional parameter were generally similar to those for the Full model (Tables 3.3 and 3.4).

In general, the scenarios investigated here were able to capture at least some of the additional mortality during 1999 and 2000, and estimated that the population has since recovered following that event to numbers that equal or possibly exceed those in 1998, but not necessarily to carrying capacity (Fig. 3.4, left and middle panels). Those scenarios which were able to fit the 1999-2000 mortality event also resulted in less precise (and slightly lower) estimates of current depletion (Fig. 3.6, right panels). There was essentially no support for carrying capacity being greater than 40,000 or that the stock size is currently at less than 70% of carrying capacity for any of the scenarios (Fig. 3.5, lower panels; Fig. 3.6, right panels; Tables 3.3 and 3.4).

The “Full” model was able to capture the variability in the calf production estimates quite well (Fig. 3.4; left panels). Even though the “Deterministic” model was not fitted to the calf estimates, the resulting estimates of calf production were consistent with the average observed calf production in recent years, albeit with much more uncertainty around these estimates than the “Full” model (Fig. 3.4; right panels). None of the scenarios considered were fitted to the 1980 and 1981 calf counts as noted above. However, there are model-predictions corresponding to those counts. In general, the predicted numbers of calves in 1980 and 1981 exceeded the observations, more so for the “Deterministic” model (although the observed values were within the 95% probability intervals for this model) (Fig. 3.4).

The “Full” model fit the data nearly equally well irrespective of whether ε_{add-1} and ε_{add-2} were estimated individually (Fig. 3.4 left and centre panels). However, the estimates of ε_{add-2} were quite different between these scenarios (Tables 3.3 and 3.4, second row last two columns). The scenario which estimated both ε_{add-1} and ε_{add-2} resulted in a larger negative value for ε_{add-2} (lower survival) and was better able to fit the strandings data during the years of the mortality event (Fig. 3.4, middle and bottom rows; Tables 3.3 and 3.4).

The parameter that related the sea-ice index to the process error deviations β was estimated to be negative for all but two scenarios (Tables 3.3 and 3.4). Setting $\sigma_i^*=0.30$ allowed the model to fit all but the most extreme years of the sea-ice index (e.g., Fig. 3.4, left and middle panels). Consequently, birth and survival rates were lower (i.e. lower calf production and higher numbers of strandings) than expected during years for which the sea-ice index was large. This result was most evident during those years before the first stranding and calf estimates. For example, calf production and survival were estimated to have been less than

otherwise expected during the heavy sea-ice years of the 1970's because of negative process error deviations during those years (Fig. 3.4).

Varying the value of σ_I^* did not greatly affect the ability of the model to fit the data other than the sea-ice index itself. The environmental signal in the process error deviations became more pronounced, especially for those years before the strandings and calf data were available, by giving the sea-ice data more weight (i.e., the $\sigma_I^*=0.10$ scenario in Tables 3.3 and 3.4). Likewise, the process error deviations were close to zero prior to there being strandings and calf data when the sea-ice data were substantially down-weighted (i.e., the $\sigma_I^*=1.00$ scenarios in Tables 3.3 and 3.4). This scenario led to a positive value for the parameter β . However, the estimated value of β for this scenario was essentially irrelevant because the process error deviates were essentially zero prior to the calf and strandings data and because of the lower weight given to the effect of sea-ice on the model fits.

There was essentially no difference between the results for the two scenarios which only fit to recent sea-ice data from 1993-2008 when the additional process error was assumed equal for birth and survival rates during the mortality event ("Recent Ice" and "Perryman *et al.* Ice" in Tables 3.3). However, the estimated values for ε_{add-1} were more negative for those scenarios which estimated both ε_{add-1} and ε_{add-2} , leading again to better fits to the strandings data and higher estimates of the numbers of animals which died during those years ("Natural Mortality '99 + '00" in Table 3.4). It followed that the estimates of current depletion were lower for those scenarios which were only fit to the recent ice data and also estimated both ε_{add-1} and ε_{add-2} (Table 3.4).

The realized standard deviation of the process error residuals (σ_ε) was similar across all stochastic scenarios and generally equal to about 0.20; with the notable

exceptions of those scenarios for which the input value for σ_ε was varied (Tables 3.3 and 3.4). Not surprisingly, varying the value for σ_ε had a direct result on the realized standard deviations of the process error residuals and, as expected, the realized standard deviations were less than the value for σ_ε . For the Full model with $\varepsilon_{add-1} = \varepsilon_{add-2}$, the point estimates for σ_ε were 0.07, 0.21 and 0.33, given $\sigma_\varepsilon = 0.30, 0.50$ and 0.70 respectively. The scenarios with $\sigma_\varepsilon = 0.30$ exhibited fairly deterministic dynamics (with the exception of the impact on the ε_{add-1} during the 1999-2000 mortality event). These scenarios also led to estimated values for $CV_{add-1} = 0.18$ and 0.20 , which were the largest of any of the scenarios considered in these analyses (Tables 3.3 and 3.4). The estimated value for β was positive when $\sigma_\varepsilon = 0.30$ and both ε_{add-1} and ε_{add-2} were estimated individually, but due to the essentially deterministic dynamics (i.e., very small process error deviates), the estimated value of β for this scenario was inconsequential. In general, the effect of increasing the standard deviation of the process errors was similar to that of decreasing the value for σ_I^* (or similarly, only fitting to the sea-ice data for 1993-2008), and vice-versa. That is, the estimated effects of sea-ice became more exaggerated for higher values of σ_ε or lower values for σ_I^* (or longer time series of sea-ice), especially during those years before calf and strandings data became available.

3.3 DISCUSSION

We incorporated an environmental index into a population dynamics modeling framework, and allowed for a hypothesized relationship between sea-ice and gray whale population dynamics when fitting to observations of abundance, sea-ice, strandings and calf production. The incorporation of such a relationship could potentially improve our understanding of cetacean population dynamics and help to

determine whether existing management strategies are robust to climate-induced forcing of the population dynamics. However, it should be noted that the goal of this investigation was not to provide evidence for or against a certain hypothesis about how environmental conditions may affect population dynamics. Rather, the primary aim was to develop a tool by which alternative hypotheses may be explicitly taken into account within a population dynamics modeling framework; ultimately, providing a means through which the robustness of management strategies may be evaluated, given such hypotheses and forecasts of future climate change.

The framework allowed for the deviations in birth and survival rates to be related to an index of sea-ice in the Bering Sea, following a plausible hypothesis about how this environmental index might be related to the population dynamics of ENP gray whales. It would be straightforward to substitute an alternative environmental index (e.g., sea-ice in the Chukchi Sea, El Niño/Southern Oscillation etc...), or some weighted combination of multiple indices into the framework developed here, but such analyses were beyond the scope of this chapter. In this study, the environmental index was used as a proxy measure for the variability in birth and survival rates, while observations of calf production and strandings numbers were taken as direct measures of the underlying variability in those life history parameters. During the years for which calf count, strandings and sea-ice data were all available, the effect of sea-ice on the population dynamics was calibrated. Then, during years when the sea-ice data were available, but prior to direct observations of calf production and strandings (i.e., during the 1960s and 1970s), the expected dynamics in birth and survival rates were extrapolated based on the values of the sea-ice index.

However, something extraordinary clearly occurred during 1999 and 2000 in terms of survival rates, as exemplified by the stranding counts (Fig. 3.4), and the ability of the different model configurations to fit the 1999-2000 mortality event

had a large influence on the results. Specifically, the results were more optimistic regarding recent stock size relative to carrying capacity for those scenarios which were not able to fit the strandings data during 1999 and 2000. Furthermore, when the strandings data and mortality event were ignored (e.g., the deterministic model), the recovery of the stock since 1930 was attributed to higher calf production during the period of increasing abundance estimates, as opposed to higher survival rates. In contrast, when the stranding data are taken into account, recovery was explained through higher survival rates and lower calf production.

It was not possible to fit the strandings data for the 1999-2000 mortality event without allowing for some additional process error in the survival rates during those years, given the assumptions made regarding the nature of the process error deviations. This additional process error could be modeled in a few different ways in addition to the approach taken in this chapter. For example, Ward *et al.* (2007) estimate the probability of an unusual or ‘catastrophic’ event by adopting a mixture distribution approach to process error. That is, they estimate if a given year was a catastrophic year, and, depending on that assessment, draw the process error for that year from one of two (regular and catastrophic) distributions (in effect we have assumed an underlying mixture distribution with a step function for 1999 and 2000, where those years are given probability of 1.0 for catastrophe, and other years assigned zero probability). An estimate of the probability of a mortality event would be of great interest when running future projections and testing the Gray Whale *SLA*, whereas the approach taken here is somewhat limited in its predictive ability of future catastrophic events. However, it is not immediately obvious that available data for ENP gray whales would be sufficient to estimate the additional parameters in a mixed model approach.

Observations of recent variability in calf production and the amount of sea-ice covering the early season feeding grounds supports the hypothesis of a relationship between sea-ice and calf production in gray whales. However, it is possible that this

relationship (if it exists) is something that has developed or strengthened within the last two decades. For example, the two early (1980s) calf production estimates are nearly equal (Fig. 3.4), yet occurred during years of disparate ice conditions (Fig. 3.2). Therefore, these early calf production data suggest some non-linearity in the relationship between population dynamics and environmental forcing. We have assumed that the relationship between calf production and sea-ice is stationary (and specifically independent of population density), and one result of this assumption was that the model predicts lower than expected calf production during much of the 1960s and 1970s, when there were higher levels of sea-ice in the Bering Sea. Consequently, the estimates of survival were higher for the stochastic than for the deterministic configurations of the model (Fig. 3.7), which allowed the stochastic model to mimic for the observed trend in the abundance data.

Modeling the interaction between population density and the effects of environmental variability on vital rates, and the probability of mortality events is beyond the scope of this chapter. However, as populations increase in density, the impact of density-independent factors on population dynamics probably become more pronounced (e.g., Durant *et al.*, 2005) and accounting for density-dependent mortality events may have implications for management strategies (e.g., Wilcox and Eldred, 2003). Therefore, if the framework presented here is used to test management strategies, a plausible set of scenarios for how such environmental affects and the probability of mortality events might change with population density should be identified. It seems unlikely that it will be possible to estimate such relationships given the amount of data available for most cetacean populations, but the approach taken here could be modified to examine different assumptions along these lines (e.g., modifying σ_l^* as a function of depletion).

The assumption that the birth and survival process errors were perfectly correlated was rather simplistic. In reality, there is likely to be some correlation, but it may be imperfect and non-linear (Eberhardt, 1977; Gaillard *et al.*, 2000). The

assumption made here was fairly inconsequential during normal stranding years, because the variability in strandings among years is generally low. However, the consequence of this assumption during the 1999-2000 mortality event was substantial. Estimating both ε_{add-1} and ε_{add-2} provided better fits to the observed strandings during this event. These scenarios illustrated the constraint placed on the ability of the model to simultaneously fit the calf and strandings data during the mortality event, when the additional process error during those years was assumed to be the same for birth and survival rates. The differences between the estimates of ε_{add-2} (roughly twice as large when estimated individually) could have implications for projections of population dynamics if future mortality events are conditioned on those estimates of ε_{add-2} .

A forecast of future sea-ice conditions will be needed to perform projections of population dynamics within this framework and hence test the Gray Whale *SLA*. Overland and Wang (2007) have provided one such forecast, based on an ensemble mean from a suite of models considered by the Intergovernmental Panel on Climate Change. We plan on including that forecast as input for population projections in the next stage of this research, using the framework presented here as an operating model while testing the Gray Whale *SLA*.

Future work may extend this framework to other cetacean stocks. Several alternative candidates exist for which a relationship between environmental conditions and cetacean population dynamics has been recognized, for example: northeast Atlantic fin whales (Lockyer, 1986); sperm whales off the Galapagos Islands (Whitehead, 1997), and; north and south Atlantic right whales (Green *et al.*, 2003; Leaper *et al.*, 2006). In addition, recent observations suggest a possible relationship between body condition and sea-ice for animals taken in the aboriginal hunt for the Bering-Chukchi-Beaufort Seas stock of bowhead whales (George *et al.*, 2009). Such information, combined with an index of calf production for this

stock (Koski *et al.*, 2008), could eventually be included in a framework similar to that presented here, and then applied to testing of the bowhead *SLA* given relevant climate forecasts.

In conclusion, the framework developed here provides a basis for testing management strategies, given a hypothesis about how environmental factors influence population dynamics and climate forecasts. The results indicated that including the environmental index and fitting to the mortality event lead to somewhat different interpretations of the population dynamics of the ENP gray whale when compared to those provided by a deterministic model. Using this framework to test the Gray Whale *SLA* should help to ensure that management is robust to a plausible range of scenarios for how future climate might impact this, and other cetacean populations.

Table 3.1

The scenarios considered in these analyses. Different data sets, parameter values and estimation techniques are outlined. The scenarios labeled ‘Deterministic’ and ‘Full’ refer to the base cases, for which a Bayesian estimation framework was adopted as well as maximum likelihood estimation (MLE). A further set of parallel runs were performed for the Full model, but not fitting the calf estimates from 1999 – 2001. The remaining scenarios are variations of the ‘Full’ base case. The ‘=’ signs denote that a parameter was not estimated but set to the value of another parameter, and ‘NA’ signifies a variable that is not applicable to a certain scenario (e.g., σ_I^* is not applicable when the sea-ice data are not considered).

Scenario	Abundance Data	Calf Data	Ice Data	Strandings data	$CV_{\text{add-1}}$ (abundance)	$CV_{\text{add-2}}$ (calf)	σ_ε
Deterministic	Yes	No	No	No	Estimated	NA	NA
Full	Yes	Yes	Yes	Yes	Estimated	$= CV_{\text{add-1}}$	0.50
$\sigma_M = 0.20$	Yes	Yes	Yes	Yes	Estimated	$= CV_{\text{add-1}}$	0.50
Recent Ice	Yes	Yes	'93-'08	Yes	Estimated	$= CV_{\text{add-1}}$	0.50
Perryman <i>et al.</i> Ice	Yes	Yes	'93-'08	Yes	Estimated	$= CV_{\text{add-1}}$	0.50
$\sigma_I^* = 0.10$	Yes	Yes	Yes	Yes	Estimated	$= CV_{\text{add-1}}$	0.50
$\sigma_I^* = 1.00$	Yes	Yes	Yes	Yes	Estimated	$= CV_{\text{add-1}}$	0.50
$\sigma_\varepsilon = 0.30$	Yes	Yes	Yes	Yes	Estimated	$= CV_{\text{add-1}}$	0.30
$\sigma_\varepsilon = 0.70$	Yes	Yes	Yes	Yes	Estimated	$= CV_{\text{add-1}}$	0.70
$CV_{\text{add-1}}$ & $CV_{\text{add-2}}$	Yes	Yes	Yes	Yes	Estimated	Estimated	0.50
No Strandings data	Yes	Yes	Yes	No	Estimated	$= CV_{\text{add-1}}$	0.50

Table 3.1 continued

Scenario	σ_M	σ_I^*	$\mathcal{E}_{add-1,t}$	$\mathcal{E}_{add-2,t}$	$\underline{\mathcal{E}}_t$	Bayesian or MLE
Deterministic	NA	NA	NA	NA	NA	Both
Full	0.10	0.30	Estimated	$= \mathcal{E}_{add-1,t}$	Estimated	Both
$\sigma_M = 0.20$	0.20	0.30	Estimated	$= \mathcal{E}_{add-1,t}$	Estimated	MLE
Recent Ice	0.10	0.30	Estimated	$= \mathcal{E}_{add-1,t}$	Estimated	MLE
Perryman <i>et al.</i> Ice	0.10	0.30	Estimated	$= \mathcal{E}_{add-1,t}$	Estimated	MLE
$\sigma_I^* = 0.10$	0.10	0.10	Estimated	$= \mathcal{E}_{add-1,t}$	Estimated	MLE
$\sigma_I^* = 1.00$	0.10	1.00	Estimated	$= \mathcal{E}_{add-1,t}$	Estimated	MLE
$\sigma_{\mathcal{E}} = 0.30$	0.10	0.30	Estimated	$= \mathcal{E}_{add-1,t}$	Estimated	MLE
$\sigma_{\mathcal{E}} = 0.70$	0.10	0.30	Estimated	$= \mathcal{E}_{add-1,t}$	Estimated	MLE
CV_{add-1} & CV_{add-2}	0.10	0.30	Estimated	$= \mathcal{E}_{add-1,t}$	Estimated	MLE
No Strandings data	NA	0.30	Estimated	$= \mathcal{E}_{add-1,t}$	Estimated	MLE

Table 3.2

The parameters and their assumed prior distributions. The abbreviations for the prior distributions include: U [uniform] and N [normal]. Footnotes below describe the sources and reasoning behind these parameter values and distributions.

Parameter	Prior distribution
Maximum non-calf survival rate, S_{1+}	U[0.950, 0.999] ^a
Maximum birth rate, b_{\max}	U[0.01, 0.99]
Difference between non-calf and calf survival, $\Delta = S_{1+} - S_0$	U[0.01, 0.25] ^b
Carrying capacity, K_{1+}	U[15 000, 70 000] ^b
Relative population size in 1930, $N_{1+,1930} / K_{1+}$	U[0.050, 0.50] ^b
Process error residuals, ε_t	N[0, σ_ε^2]
Additional process error during '99/'00 mortality event, ε_{add-1} and ε_{add-2}	U[-4.0, 4.0] ^b
Influence of sea-ice on calf production, β	U[-400, 400] ^b

a. Equal to the prior distribution used in recent assessments (IWC, 1998).

b. Preliminary analyses provided no evidence of posterior support for values outside this range.

Table 3.3

Results for the scenarios based on the Full model. The estimates correspond to the mode of the posteriors for the Bayesian analyses, and MLEs are shown for the ML analyses: \tilde{S} is the median stochastic survival rate over all years; σ_{ϵ} is the standard deviation of the realized process errors (after the transformation in Eqns. 3.2 and 3.3), and the combined natural mortality during 1999-2000, corresponding to the total number of whales estimated to have died during the mortality event is also shown. The asterisk on $CV_{\text{add-2}}$ (calf) indicates that this parameter was estimated to be zero.

Scenario	$N_{1+}, 2009 / K_{1+}$	K_{1+}	b_{max}	S_{1+}	\tilde{S}_{1+}^*	S_0	\tilde{S}_0^*	σ_{ϵ}
Deterministic	0.981	22,621	0.990	0.964	NA	0.714	NA	NA
Full	0.979	26,773	0.32	0.989	0.981	0.979	0.971	0.21
$\sigma_M = 0.20$	0.996	23,159	0.65	0.978	0.971	0.968	0.961	0.22
Recent Ice	0.938	26,650	0.31	0.989	0.982	0.979	0.972	0.18
Perryman <i>et al.</i> Ice	0.940	26,054	0.33	0.988	0.980	0.977	0.969	0.18
$\sigma_I^* = 0.10$	0.980	29,784	0.25	0.995	0.989	0.960	0.954	0.23
$\sigma_I^* = 1.00$	0.941	26,162	0.33	0.988	0.980	0.978	0.970	0.18
$\sigma_{\epsilon} = 0.30$	0.941	28,889	0.28	0.994	0.985	0.984	0.975	0.07
$\sigma_{\epsilon} = 0.70$	0.962	30,890	0.21	0.997	0.993	0.870	0.866	0.33
$CV_{\text{add-1}}$ & $CV_{\text{add-2}}$	0.988	26,578	0.36	0.987	0.980	0.977	0.970	0.22
No Strandings data	0.992	22,454	0.99	0.980	0.972	0.742	0.735	0.23

Table 3.3 continued

Scenario	β	Natural Mortality '99 + '00	$CV_{\text{add-1}}$ (abundance)	$CV_{\text{add-2}}$ (calf)	$\mathcal{E}_{\text{add-1},t}$ (birth)	$\mathcal{E}_{\text{add-2},t}$ (survival)
Deterministic	NA	2,414	0.11	NA	NA	NA
Full	-2.32	2,455	0.12	0.12	-1.01	-1.01
$\sigma_M = 0.20$	-2.25	2,072	0.09	0.09	-0.44	-0.44
Recent Ice	-2.28	2,575	0.11	0.11	-0.91	-0.91
Perryman <i>et al.</i> Ice	-2.11	2,679	0.09	0.09	-0.66	-0.66
$\sigma_I^* = 0.10$	-3.20	1,996	0.19	0.19	-1.20	-1.20
$\sigma_I^* = 1.00$	-0.93	2,688	0.10	0.10	-1.05	-1.05
$\sigma_{\mathcal{E}} = 0.30$	-2.97	2,538	0.18	0.18	-1.19	-1.19
$\sigma_{\mathcal{E}} = 0.70$	-1.85	1,314	0.14	0.14	-1.22	-1.22
$CV_{\text{add-1}}$ & $CV_{\text{add-2}}$	-2.30	2,351	0.13	0.00*	-0.87	-0.87
No Strandings data	-2.14	1,468	0.08	0.08	-0.06	-0.06

Table 3.4

As for table 3.3, except that ε_{add-1} & ε_{add-2} are estimated individually. That is, the 1999/2000 event was not assumed to have an identical impact on birth and survival rates.

Scenario	$N_{1+, 2009} / K_{1+}$	K_{1+}	b_{\max}	S_{1+}	\tilde{S}_{1+}^*	S_0	\tilde{S}_0^*	σ_{ε}
Deterministic	0.981	22,621	0.990	0.964	NA	0.714	NA	NA
Full (ε_{add-1} & ε_{add-2})	0.940	29,632	0.22	0.999	0.995	0.773	0.770	0.23
$\sigma_M = 0.20$	0.996	22,960	0.705	0.978	0.971	0.968	0.961	0.22
Recent Ice	0.920	27,985	0.268	0.995	0.989	0.869	0.863	0.18
Perryman <i>et al.</i> Ice	0.911	27,604	0.277	0.996	0.990	0.803	0.798	0.19
$\sigma_I^* = 0.10$	0.959	30,835	0.207	0.999	0.995	0.849	0.845	0.24
$\sigma_I^* = 1.00$	0.711	30,711	0.230	0.996	0.991	0.746	0.741	0.21
$\sigma_{\varepsilon} = 0.30$	0.884	31,209	0.226	0.999	0.995	0.809	0.804	0.05
$\sigma_{\varepsilon} = 0.70$	0.953	29,535	0.213	0.999	0.996	0.749	0.746	0.37
CV_{add-1} & CV_{add-2}	0.954	29,799	0.233	0.999	0.995	0.773	0.770	0.24
No Strandings data	0.996	22,960	0.705	0.978	0.971	0.968	0.961	0.22

Table 3.4 continued

Scenario	β	Natural Mortality '99 + '00	$CV_{\text{add-1}}$ (abundance)	$CV_{\text{add-2}}$ (calf)	$\mathcal{E}_{\text{add-1},t}$ (birth)	$\mathcal{E}_{\text{add-2},t}$ (survival)
Deterministic	NA	2,414	0.11	NA	NA	NA
Full ($\mathcal{E}_{\text{add-1}}$ & $\mathcal{E}_{\text{add-2}}$)	-2.25	2,620	0.12	0.12	-0.99	-2.24
$\sigma_M = 0.20$	-2.23	2,314	0.09	0.09	-0.56	0.15
Recent Ice	-2.31	2,981	0.11	0.11	-1.45	-0.85
Perryman <i>et al.</i> Ice	-2.06	3,414	0.09	0.09	-1.49	-0.55
$\sigma_I^* = 0.10$	-3.08	2,040	0.19	0.19	-1.92	-1.07
$\sigma_I^* = 1.00$	0.94	5,187	0.09	0.09	-2.39	-1.26
$\sigma_{\mathcal{E}} = 0.30$	3.07	2,854	0.20	0.20	-2.08	-1.19
$\sigma_{\mathcal{E}} = 0.70$	-1.75	2,359	0.13	0.13	-2.38	-0.94
$CV_{\text{add-1}}$ & $CV_{\text{add-2}}$	-2.23	2,549	0.14	0.00*	-2.20	-0.95
No Strandings data	-2.23	2,314	0.09	0.09	-0.56	0.15

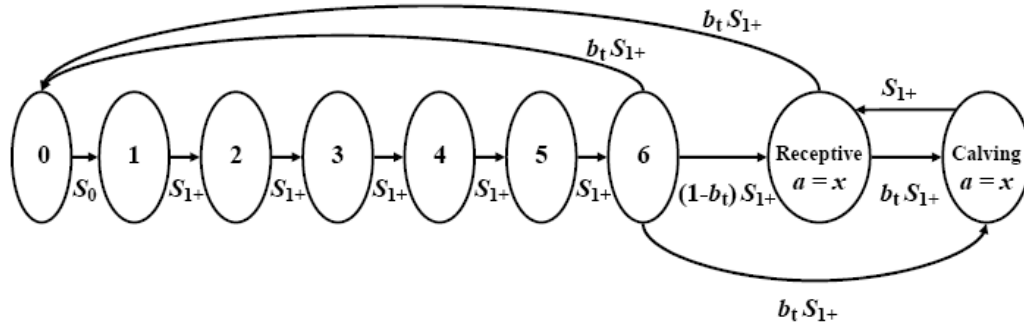


Figure 3.1: Life cycle graph of the model used to track the number of females in each reproductive stage through time. This life cycle refers to the underlying deterministic model, with transition probabilities shown as functions of life history parameters. The survival and birth rates were modified to be stochastic in all of the analyses presented here (except ‘Deterministic’). The arrow from immature to calf arises because some juveniles may mature and give birth (i.e. become pregnant at first estrous) during the projection interval from time t to $t+1$.

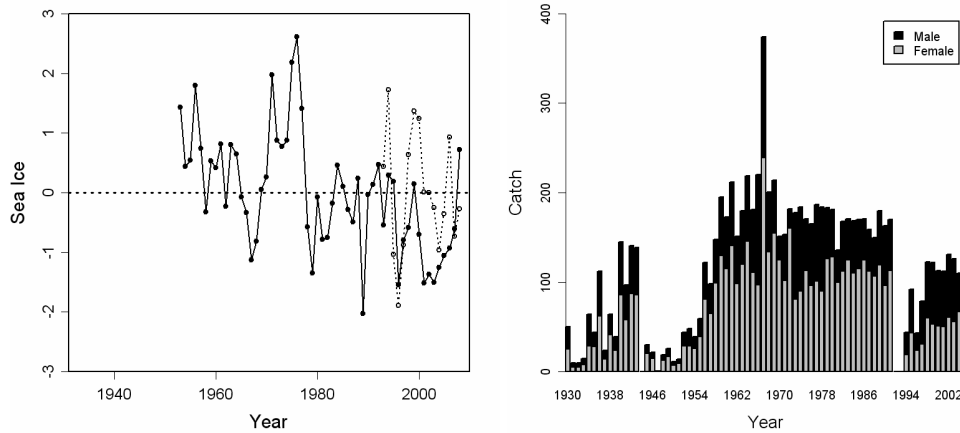


Figure 3.2: (Left panel) The standardized HadSST index for the March-April averaged sea-ice area covering the Bering Sea is shown by the solid line, and the Perryman *et al.* index is shown as the dashed line. Positive values represent years with greater than average spring ice over the time period considered. (Right panel) Catches by individuals and sex: 1930-2006.

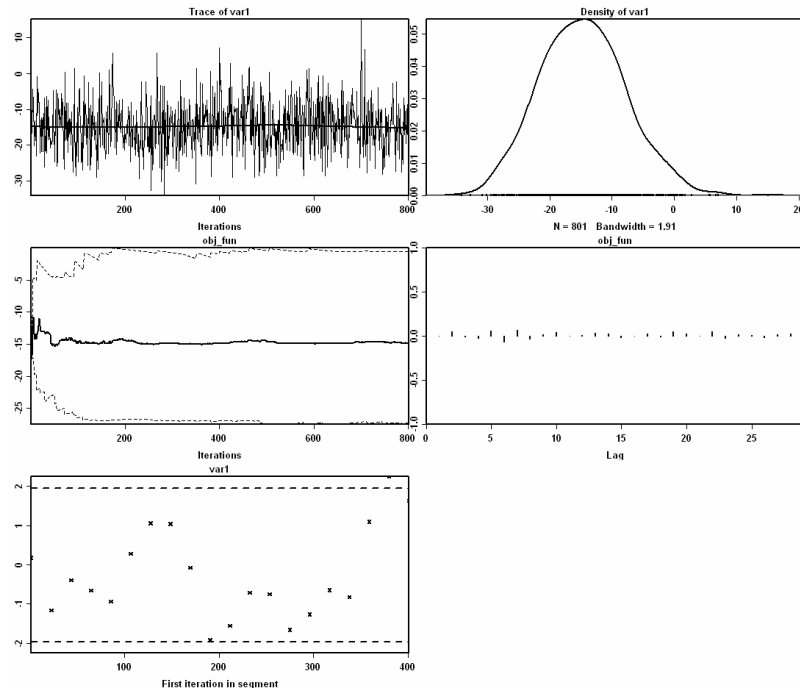


Figure 3.3: Diagnostic plots for the negative log-likelihood function resulting from the MCMC chain for the base-case Full model scenario (all data). Clockwise from upper left: trace, density, autocorrelation, Geweke's z-score, and the cumulative quantile plots showing the evolution of the median (solid line) and 95th percentiles of the chain.

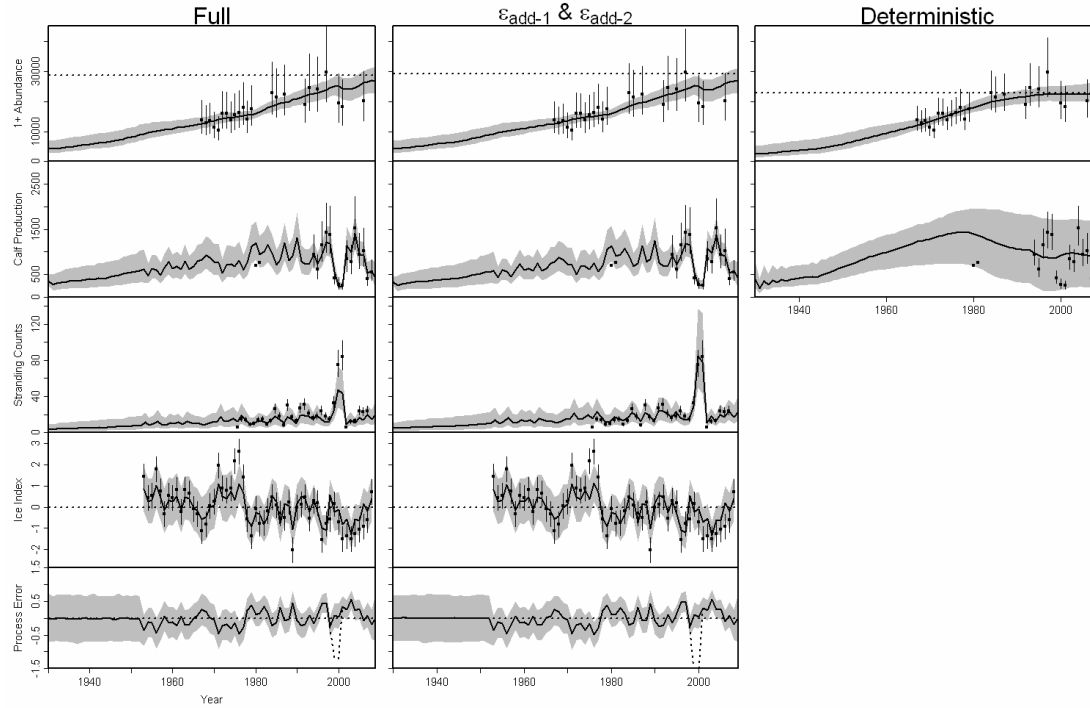


Figure 3.4: Model fits are shown: left column is the Full model; middle column is that model, but estimating ε_{add-1} and ε_{add-2} separately, and; right column is the Deterministic model. From top to bottom: abundance; calf production; strandings; sea-ice, and; estimated process error deviations. The abundance estimates are plotted with the 95% CIs associated with the mode of the posterior distribution for CV_{add-1} . The median of the posterior estimate for carrying capacity is plotted as a horizontal line with the abundance fits. For all plots, the medians and 95% Bayesian credibility intervals are shown as solid lines and shaded areas respectively. The calf estimates for 1980-81 were not fit for the first two scenarios, nor were any of the recent calf estimates fit for the Deterministic model. However, they are plotted for reference. Horizontal dotted lines at zero are plotted in the fits to the sea-ice data and the process error deviation estimates for reference. And the median of the posterior for ε_{add-2} is represented by the more vertical dashed line on the bottom two plots.

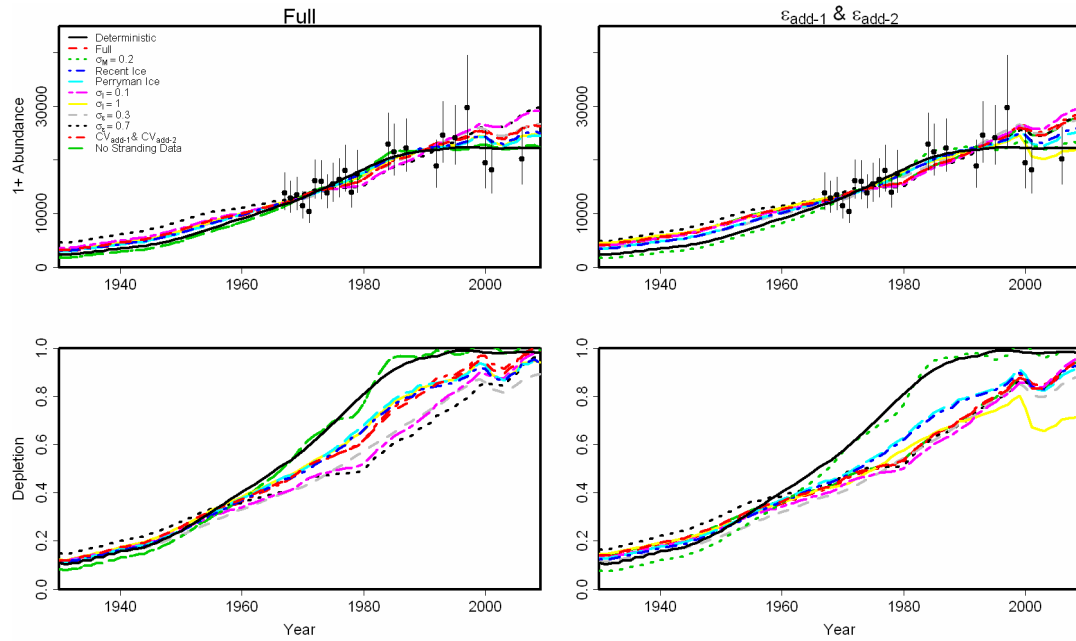


Figure 3.5: Fits to the abundance data based on the maximum likelihood estimates for each scenario (upper panels) and the estimated relative population size (“depletion”) through time (bottom panels). The scenarios under the Full model are shown in the left panels, while those that estimated ε_{add-1} and ε_{add-2} are shown in the right panels. The deterministic scenario is plotted as the solid black line. 95% CIs are plotted for the abundance estimates assuming a value for $CV_{add-1} = 0.10$.

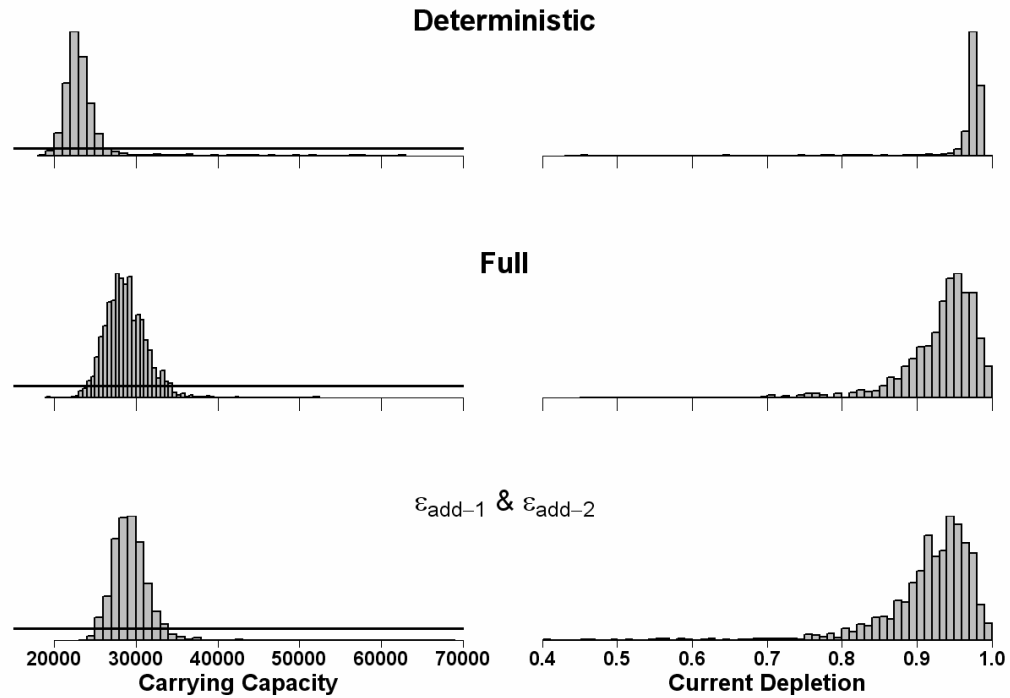


Figure 3.6: Histograms comparing marginal posterior densities (bars). Estimates of carrying capacity (plots on left side) and current depletion (right side) are shown. The upper row shows the samples from the posterior for the deterministic model, the middle rows shows samples from the posteriors for the Full model and the bottom row shows those for the Full model when ε_{add-1} and ε_{add-2} are estimated individually. The uniform prior for carrying capacity is shown as a solid line.

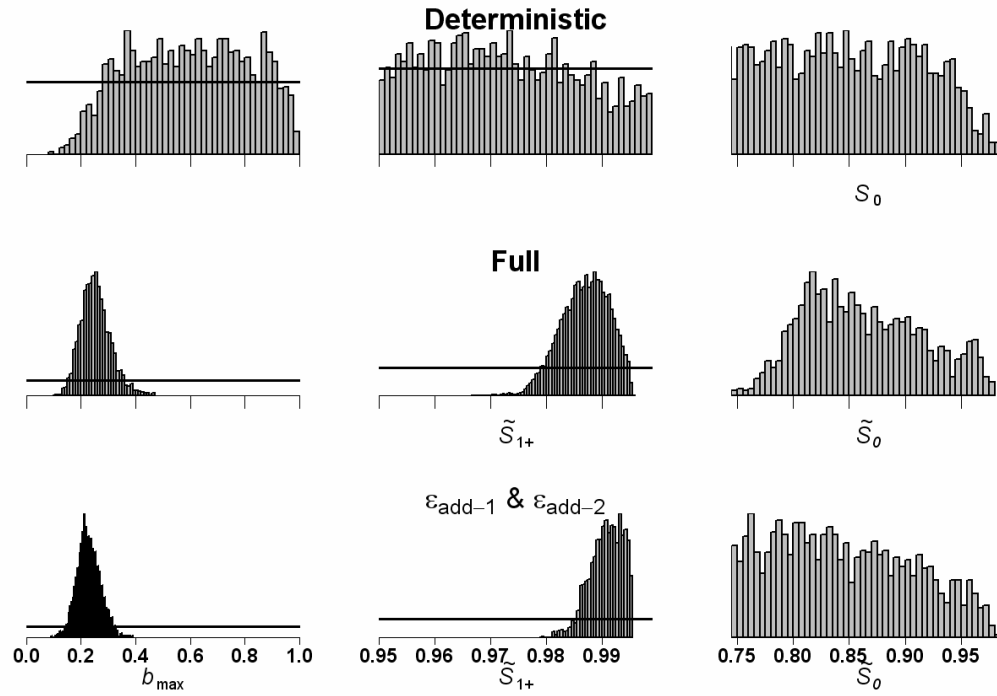


Figure 3.7: Histograms comparing marginal posterior densities (bars). Estimates of the maximum birth rate (in the limit of vanishing population size) (left column), survival rate of non-calves S_{1+} (middle column), and calf survival S_0 (right column) are shown. The upper row shows the samples from the posterior for the deterministic model, the middle rows shows samples from the posteriors for the Full model and the bottom row shows those for the Full model when ε_{add-1} and ε_{add-2} are estimated individually. The survival rates for the Full model (including those for the ε_{add-1} and ε_{add-2} scenario) are shown as the medians through time, in order to provide a better comparison with those estimates from the deterministic model. The uniform priors are shown as solid lines.

Chapter 4:

Testing the Gray Whale *SLA*: allowing environmental variability to influence population dynamics**ABSTRACT**

The performance of the Gray Whale *SLA* was evaluated based on an operating model which was conditioned on available information, including survey estimates of 1+ abundance, calf counts, strandings data, and the extent of sea-ice in the early season feeding grounds in the Bering Sea. The scenarios considered in the analyses explore the impact of different sources of environmental variation, including scenarios in which future environmental forcing and episodic events are driven by the relationship between extent of sea-ice and reproductive success and survival. A variety of sources of uncertainty are considered, including parameter uncertainty, the uncertainty about the relationship between the extent of sea-ice and population dynamics, and observation error. The impact of these sources of uncertainty on the performance of the Gray Whale *SLA* appears small.

4.1 INTRODUCTION

The eastern North Pacific (ENP) stock of gray whales is currently subject to aboriginal hunting, with recommended strike limits based on the *Gray Whale Strike Limit Algorithm (SLA)* under the Aboriginal Subsistence Whaling Management Procedure (AWMP) of the IWC (IWC, 2003). *Implementation Reviews* are scheduled under the AWMP every five years, and that for the Gray Whale *SLA* is currently due. The goal of *Implementation Reviews* is to evaluate new information that has become available since the last *Implementation Review* (or the original *Implementation*) and to determine whether the current state of nature is not outside the realm of plausibility envisioned during the testing of the original *SLA*. If this is the case, additional simulation trials may be conducted to assess whether the

performance of the adopted *SLA* remains reasonable, and if not, what changes to the *SLA* are needed.

New or updated sources of information pertaining to the population dynamics of ENP gray whales have become available in recent years and need to be considered during this *Implementation Review*, including: (1) new abundance estimates (Rugh *et al.*, 2008); (2) new estimates of calf production during 1994-2008 from the northbound migration at Point Piedras Blancas, California (Perryman *et al.*, 2002; Perryman, *unpublished data*), and; (3) the number of stranded animals on the coasts of California, Oregon and Washington states, for which a combined annual count is available for 1975-2006 (Brownell Jr. *et al.*, 2007). The latter potentially contains information on the magnitude of the mortality event during 1999/2000 (Gulland *et al.*, 2005). In addition to these data sets, it has been hypothesized that observed variability in the calf counts is a function of the amount of sea-ice covering the early season feeding grounds (Perryman *et al.*, 2002).

Therefore, in this chapter we test the performance of the *SLA* given scenarios for which future population dynamics are subject to environmental forcing and episodic events, using an operating model that integrates these sources of new information and the hypothesis of environmental forcing on the population dynamics (Chapter 3). A forecast of relevant sea-ice conditions based on global climate model output (Overland and Wang, 2007) is used to modify the future stochastic birth and survival rates when testing the *SLA*, given the estimated relationship between observed variations in recent sea-ice (Rayner *et al.*, 2003) and calf and strandings data. This approach involves the incorporation of climate-model based forecasts into the operating model; the same basic approach is also being used to test the performance of alternative management strategies in other fisheries (e.g., Gulf of Alaska walleye Pollock, *Theragra chalcogramma*; A'mar *et al.*, 2009).

Standard summary statistics are provided for the trials investigated here, and these are compared to recent results from the *Evaluation Trials* provided by Punt and Breiwick (2008) to the extent possible. The analyses presented here should help to ensure that the performance of the current *SLA* remains satisfactory (or else provide insight into potential weaknesses), given the new information that has become available since the previous phase of testing and adoption (IWC, 2005a).

4.2 METHODS

Operating model

The population dynamics model developed in Chapter 3 (corresponding to the ‘Full’ scenario) was used as the operating model. This model is sex- and age-based, with an annual time-step. The dynamics includes stochastic birth and survival rates, and explicitly considers the transition between receptive and calving stages for mature females (Fig. 4.1). For consistency, the notation of Chapter 3 is adopted below.

Density dependence was assumed to act through the birth rate according to a Pella-Tomlinson function of 1+ depletion:

$$b_t = \max \left\{ 0, b_{eq} + (b_{max} - b_{eq}) \left[1 - \left(\frac{N_{1+,t}}{K_{1+}} \right)^z \right] \right\} \quad (4.1)$$

where:

b_{max} is the maximum birth rate (in the limit of zero population size);

K_{1+} is the carrying capacity of the 1+ component of the population (all animals aged 1 yr and older);

b_{eq} is the equilibrium birth rate at carrying capacity;

- z is the degree of density-dependent compensation (assumed to equal 2.39, which implies maximum sustainable yield at a population size approximately 60% of K_{1+}), and;
- $N_{1+,t}$ is the size of the 1+ component of the population (both sexes combined) in year t .

Selectivity was assumed to be knife-edged and uniform on ages 5+, and the population trajectories were initialized in 1930.

The operating model was conditioned on available data, including: (1) estimates of population size during 1967-2006 (starting year of survey) from the southbound migration at Granite Canyon, California (Rugh *et al.*, 2005, 2008); (2) estimates of calf production during 1994-2008¹³ from the northbound migration at Point Piedras Blancas, California (Perryman *et al.*, 2002; Perryman, *unpublished data*), (3) the number of stranded animals on the coasts of California, Oregon and Washington states, for which a combined annual count is available for 1975-2006 (Brownell Jr. *et al.*, 2007)¹⁴, and; (4) estimated sea-ice area covering the Bering Sea, averaged over March and April during 1953-2008, as calculated by the Hadley Center for their sea ice and sea surface temperature data set version 1 ('HadSST') (Rayner *et al.*, 2003) (Fig. 4.2, left panel).

Deviations from expected birth and survival rates were allowed to be a function of sea-ice variability in the Bering Sea. Thus, the model is an adaptation of the hypothesis that the amount of sea-ice in the Bering Sea early during the feeding season may be related to variability in calf production the following year (Perryman *et al.*, 2002).

¹³ The two early estimates of calf production during 1980-1981 (Poole, 1984) were not used in these analyses.

¹⁴ Data on strandings are collected in other locations (e.g. Mexico and Alaska), but the stranding network effort in California, Oregon and Washington has been more consistent through the years.

Future projections

The population was projected forward from the start of 2009. Values for the environmental index were based on an ensemble mean forecast of future sea-ice in the Bering Sea (March-April average) (Overland and Wang, 2007). The trials were based on a 92-year time horizon ($T=92$), because the time series of forecasted sea-ice was only available through 2008. Each simulated trajectory was based on a set of parameter values $\underline{\theta}_i$ (e.g., K_{1+} , b_{\max} etc...) sampled from the joint Bayesian posterior distribution constructed using the MCMC algorithm described in Chapter 3. In a given year, the process error residuals about the expected birth and survival rates were:

$$\varepsilon_t = \left(I_t^{obs} / \beta \right) - \gamma_t \quad (4.2)$$

where:

- I_t^{obs} is the forecasted value of the environmental index for year t (Fig. 4.2, left panel);
- β is a scaling parameter that accounts for the influence of the environment on the process error residuals (sampled from the joint posterior);
- γ_t is a generated normal random deviate reflecting error about the sea-ice – process error relationship, such that $\gamma_t \sim N(0; \sigma_I^2)$, and;
- σ_I is the standard deviation of the residual error for the environmental index:

$$\sigma_I = |\beta| \sigma_I^* \quad (4.3)$$

This formulation takes a fixed input value for σ_I^* (assumed to be 0.30 for these analyses, corresponding with the ‘Full’ model described in Chapter 3) and scales

the expected standard deviation of the fits to the environmental index by the absolute value sampled from the posterior distribution for β .

Stochastic birth and survival rates

The stochastic survival and birth rates were calculated given the generated process errors for each year. Birth rates were assumed to vary annually about the deterministic value given by Eqn. 4.1. Since this rate must lie between zero and one, its realization in any one year was calculated using a logistic transformation:

$$b_t^* = \left[1 + \exp(-(\Phi^{-1}(b_t) \sqrt{2.76 + \sigma_\varepsilon^2} + \varepsilon_t + \varepsilon_{add-1,t})) \right]^{-1} \quad (4.4)$$

where:

Φ^{-1} is the inverse standard normal cumulative distribution function;

ε_t is the process error deviation for year t , and;

$\varepsilon_{add-1,t}$ allows for additional process error in the birth rate during years with extraordinary dynamics, such as 1999 and 2000 (in other years, this parameter was set equal to zero).

This formulation of stochastic birth rates ensured that the expected birth rate in a given year was equal to the deterministic value from Eqn. 4.1.

Survival rates were also allowed to vary annually with the same process error residuals as birth rates. It was assumed that these rates were independent of sex and perfectly correlated between ages in a given year, so that:

$$S_{a,t}^* = \left[1 + \exp(-(\Phi^{-1}(S_a) \sqrt{2.76 + \sigma_\varepsilon^2} + \varepsilon_t + \varepsilon_{add-2,t})) \right]^{-1} \quad (4.5)$$

where:

$S_{a,t}^*$ is the realized age-specific survival rate during year t ;

S_a is the expected survival rate from age a to age $a+1$; and

$\varepsilon_{add-2,t}$ is a parameter which allows for additional process error in survival rates during years with extraordinary dynamics, such as 1999 and 2000 (in other years, this parameter was set equal to zero).

For these analyses, the additional process error in survival rates was assumed to be equal to that for birth rates (i.e., $\varepsilon_{add-1,t} = \varepsilon_{add-2,t} = \varepsilon_{add,t}$).

Data generation

Future abundance estimates were assumed to become available every 10 years. Observation error was assumed to be log-normal:

$$N_{1+,t}^{obs} = N_{1+,t} e^{\phi_t} \quad (4.6)$$

where:

$N_{1+,t}^{obs}$ is the survey estimate of 1+ abundance for year t ;

$N_{1+,t}$ is the ‘true’ 1+ abundance at the start of year t ;

ϕ_t is a normal random deviate $\sim N(0, \sigma^2)$; where $\sigma = \sqrt{\overline{CV}_{est}^2 + CV_{add-1}^2}$;

CV_{add-1} is the extent of additional error about the abundance estimates (sampled from the joint posterior), and;

\overline{CV}_{est} is the expected (sampling) standard deviation of the logarithm of $N_{1+,t}^{obs}$:

$$\overline{CV}_{est} = \sqrt{\frac{1}{Y} \sum_{y=1}^Y CV_y^2}, \quad (4.7)$$

where:

y indexes years for which there are survey data up to 2008, and;

Y is the total number of such years.

The estimates of abundance and \overline{CV}_{est} (as opposed to σ) were passed to the *SLA*. No attempt was made to account for further estimation error in the abundance estimates (i.e., mean school size estimation error calculations were ignored).

Need

The annual need Q_t for year t was calculated according to the ‘need envelope’:

$$Q_t = Q_{2009} + \frac{t - 2009}{91}(Q_{2098} - Q_{2009}) \quad (4.8)$$

where:

Q_{2009} (=150) is the present need, and;

Q_{2098} is the final need (in year 2098).

The level of need supplied to the *SLA* was the total (block) need for the 5-year period for which the strike limits were to be set. Two values were assumed for final need (in yr. 2098), corresponding with the ‘base case’ ($Q_{2098}=340$) and ‘high need’ ($Q_{2098}=530$) trial levels used in previous testing of the *SLA* (IWC, 2003).

Trials

The set of trials is listed in Table 4.1. In addition to the two levels of final need, six scenarios were explored with respect to the future probability (if any) of catastrophic (otherwise known as ‘episodic’) events and the nature of stochastic (or deterministic) population dynamics: (H0) Deterministic population dynamics with no future catastrophic events¹⁵; (H1) Environmental stochasticity (as a function of sea-ice) with no future catastrophic events; (H2) Environmental stochasticity (as a function of sea-ice) with probability of future catastrophic events conditioned on the stranding index (corresponding to the percentage of catastrophic years¹⁶ during the time series of stranding counts); (H3) Environmental stochasticity (as a function

¹⁵ The two deterministic trials are most comparable with the base-case operating models in IWC (2004).

¹⁶ The 2 years (1999 and 2000) during the unusual mortality event were considered to be catastrophic.

of sea-ice) with the probability of future catastrophic events p^* conditioned on the percentage of times they occurred during the fitting process when 1+ depletion was greater than 0.40 (Eqn. 4.9; Fig. 4.2 right); (H4) As for H3, but the environmental stochasticity was independent of the sea-ice index, i.e. simply $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$, and; (H5) As per H4 but with no future catastrophes.

A depletion of 0.4 represents a level encompassing the full range of trajectories from the posterior (i.e., a small number of those trajectories were estimated to have never recovered to more than 50% of carrying capacity, when carrying capacity was estimated to have been around 50,000 individuals). The probability of future catastrophes p^* conditioned on the percentage of times they occurred during the fitting process when 1+ depletion was greater than 0.40 was then:

$$p^* = 2 \left[\left(\sum_{t=1930}^{2008} I(N_{1+,t} / K > 0) \right)^{-1} \right] \quad (4.9)$$

where:

$I()$ is the indicator function.

Performance statistics

The performance statistics were calculated based on future block quotas returned from the standalone version of the ‘GUP2’ *SLA* (Punt and Breiwick, 2008). All performance statistics were computed in terms of the 1+ component of the population following the standard methods and notation of the AWMP (IWC, 2003). Specifically, four performance statistics were calculated:

1. (D1) Final depletion: $N_{1+,2098} / K_{1+}$;
2. (D8) Rescaled final population size: $N_{1+,2098} / N_{1+,2098}^*$;

where:

$N_{1+,2098}^*$ is the 1+ population size in the final year T , under a scenario of zero future catches.

3. (D10) Relative increase: $N_{1+,2098} / N_{1+,2009}$, and;

4. (N9) Average need satisfaction: $\frac{1}{T} \sum_{t=2009}^{2098} \frac{C_t}{Q_t}$.

where:

C_t is the catch during year t , which is determined by the *SLA* through the 5-year block quota system.

4.3 RESULTS

1601 simulations were run for each scenario, corresponding to the number of samples from the posterior provided by the analyses of Chapter 3. In general, the Gray Whale *SLA* was able to satisfy need and maintain a population size near carrying capacity for each of scenarios examined in these analyses. For example, all of the scenarios with base need had an average need satisfaction of 100% and the lowest median final 1+ depletion was 0.874 (Table 4.2). Not surprisingly, those scenarios with higher final need resulted in lower final depletion levels and lower average need satisfaction. However, the differences were not very large (e.g., the lowest median 1+ depletion for the high need scenarios was 0.817). Moreover, none of the scenarios resulted in a lower 5th percentile for the final 1+ depletion less than 0.60. The relative increase statistic (D10) was close to 1 for all scenarios, which indicates stability in the population dynamics. This is, however, not unexpected given the results of Chapter 3, which suggest that this population is close to carrying capacity at present.

The annual probability of future catastrophes for the two ‘H2’ scenarios was 0.0625, as determined by the number of years for which an episodic event was observed, divided by the total number of years in the strandings index

(2yrs/32yrs)(Brownell *et al.*, 2007). The distribution of probabilities of future catastrophes for the ‘H3’ and ‘H4’ scenarios is shown in Fig. 4.2 (right panel). The probability of future catastrophe ranged between 0.025 and 0.222 for those scenarios, with a median of 0.043, which was less than that when conditioned on the stranding index. However, the average difference between these two approaches was relatively small, as evidenced by the nearly identical results between these two assumptions (Table 4.2; Fig. 4.3).

The predicted area of sea-ice on the Bering Sea feeding grounds is forecasted to decrease dramatically, with less than 50% of the average observed area of sea-ice in March-April during future decades (Fig. 4.2, left panel) (Overland and Wang, 2007). The scenarios (H1, H2, and H3) with population dynamics that were a function of this environmental index resulted in the most optimistic outcomes (Table 4.2), with some final depletion levels which were slightly greater than 1.0. On the other hand, the two scenarios that modeled generic environmental stochasticity independent of sea-ice (H4 and H5), resulted in the most pessimistic final depletion levels of any of the scenarios investigated (Table 4.2). Likewise, the trend in process error residuals was very different between these two sets of scenarios. Those scenarios which modeled process error as a function of future sea-ice resulted in an increasing trend in process error deviations, while those scenarios which modeled environmental stochasticity as an independent process resulted in no such trend (Fig. 4.4). However, in terms of the median average need satisfaction, there was essentially no difference between any of the scenarios (Table 4.2).

The results of the “deterministic” trials (H0) were more optimistic than those of the corresponding trials on which the Gray Whale *SLA* was based (GE01 and GE14) (compare table 2 of Punt and Breiwick (2008) with the results for the two H0 trials in table 4.2 of this chapter). However, the differences in the values for the performance statistics are slight, and qualitatively the results of trial H0 and GE01

are identical. The differences in results are attributable to a variety of causes, including differences in the population dynamics models, in the data used to condition the operating model, and in the priors for the parameters of the model.

4.4 DISCUSSION

The analyses incorporated an index of sea-ice variability into an operating model which was used to test the Gray Whale *SLA*, given forecasts for future climate change and a hypothesis regarding the interaction between sea-ice and population dynamics. The trials presented here differ from the standard set designed by the Standing Working Group of the AWMP, in that they were explicitly conditioned on the most recently available data and a hypothesis regarding environmental forcing. For example, deviations in the survival rates during the 1999/2000 mortality event (and resulting population sizes at the start of the future trajectories) were conditioned on observed variability in the strandings data. A set of several alternative trials was also preformed, to compare the results of the environmental forcing scenario to those for which future population dynamics were assumed to be deterministic, or to be subject to random environmental stochasticity (i.e., ignoring sea-ice). For all of the scenarios considered here, the Gray Whale *SLA* was able to maintain stock size and satisfy need at very high levels. Therefore, there is no indication from these analyses that any revisions to the *SLA* are necessary at this time.

It is interesting to note that the assumption that the population dynamics were related to sea-ice led to more optimistic results. This was essentially the result of extrapolating (based on those years for which calf production and strandings data exist) a recent relationship between the environment and population dynamics into the future, under the assumption that such an effect (if it exists) would be constant with respect to time and population density (among other factors). While this is obviously an oversimplification, the framework used here could be modified during

the next *Implementation Review* in order to take into account alternative hypotheses with respect to predicted changes in the effect of future environmental variability on population dynamics (e.g., by modifying σ_I^* as a function of depletion).

Table 4.1

The scenarios considered. The trials are denoted by an ‘H’ followed with the trial number and then ‘BN’ or ‘HN’ for base or high final need. Descriptions are given for each scenario in terms of the stochastic or deterministic nature of the population dynamics and the probability of future catastrophes. The extent of future stochasticity σ_ε is equal to 0.50 (for consistency with the ‘Full’ analyses of Chapter 3) for all except the deterministic scenario.

Trial	Description	σ_ε	Final need	Probability of future catastrophe	Future stochasticity
H0 : BN	Deterministic + no future catastrophes	NA	340	0	None (Deterministic)
H1 : BN	Environmental stochasticity + no future catastrophes	0.5	340	0	Environmental
H2 : BN	Environmental stochasticity + p(future catastrophe)= 0.0625	0.5	340	0.0625	Environmental
H3 : BN	Environmental stochasticity + p(future catastrophe)= p^*	0.5	340	p^* (Eqn. 4.9)	Environmental
H4 : BN	Stochasticity (no sea-ice) + p(future catastrophe)= p^*	0.5	340	p^* (Eqn. 4.9)	Environmental (no sea-ice)
H5 : BN	Stochasticity (no sea-ice) + no future catastrophes	0.5	340	0	Environmental (no sea-ice)
H0 : HN	Deterministic + no future catastrophes	NA	530	0	None (Deterministic)
H1 : HN	Environmental stochasticity + no future catastrophes	0.5	530	0	Environmental
H2 : HN	Environmental stochasticity + p(future catastrophe)= 0.0625	0.5	530	0.0625	Environmental
H3 : HN	Environmental stochasticity + p(future catastrophe)= p^*	0.5	530	p^* (Eqn. 4.9)	Environmental
H4 : HN	Stochasticity (no sea-ice) + p(future catastrophe)= p^*	0.5	530	p^* (Eqn. 4.9)	Environmental (no sea-ice)
H5 : HN	Stochasticity (no sea-ice) + no future catastrophes	0.5	530	0	Environmental (no sea-ice)

Table 4.2

The medians, and upper and lower 5th percentiles of the performance statistics for each scenario. See text for the definitions for each of the performance statistics

Trial	Description	D1: Final 1+ Depletion		
		5%	Median	95%
H0 : BN	Deterministic + no future catastrophes	0.908	0.933	0.950
H1 : BN	Environmental stochasticity + no future catastrophes	0.940	0.981	1.030
H2 : BN	Environmental stochasticity + p(future catastrophe)= 0.0625	0.914	0.974	1.026
H3 : BN	Environmental stochasticity + p(future catastrophe) = p^*	0.922	0.976	1.027
H4 : BN	Stochasticity (no sea-ice) + P(future catastrophe) = p^*	0.745	0.874	0.953
H5 : BN	Stochasticity (no sea-ice) + no future catastrophes	0.802	0.897	0.960
H0 : HN	Deterministic + no future catastrophes	0.855	0.899	0.927
H1 : HN	Environmental stochasticity + no future catastrophes	0.913	0.963	1.017
H2 : HN	Environmental stochasticity + p(future catastrophe)= 0.0625	0.880	0.954	1.011
H3 : HN	Environmental stochasticity + p(future catastrophe) = p^*	0.894	0.957	1.013
H4 : HN	Stochasticity (no sea-ice) + p(future catastrophe) = p^*	0.657	0.817	0.917
H5 : HN	Stochasticity (no sea-ice) + no future catastrophes	0.722	0.847	0.927

Table 4.2 continued

Trial	Description	D8: Rescaled 1+ Depletion		
		5%	5%	5%
H0 : BN	Deterministic + no future catastrophes	0.875	0.875	0.875
H1 : BN	Environmental stochasticity + no future catastrophes	0.910	0.910	0.910
H2 : BN	Environmental stochasticity + $p(\text{future catastrophe}) = 0.0625$	0.886	0.886	0.886
H3 : BN	Environmental stochasticity + $p(\text{future catastrophe}) = p^*$	0.896	0.896	0.896
H4 : BN	Stochasticity (no sea-ice) + $P(\text{future catastrophe}) = p^*$	0.731	0.731	0.731
H5 : BN	Stochasticity (no sea-ice) + no future catastrophes	0.775	0.775	0.775
H0 : HN	Deterministic + no future catastrophes	0.833	0.833	0.833
H1 : HN	Environmental stochasticity + no future catastrophes	0.889	0.889	0.889
H2 : HN	Environmental stochasticity + $p(\text{future catastrophe}) = 0.0625$	0.858	0.858	0.858
H3 : HN	Environmental stochasticity + $p(\text{future catastrophe}) = p^*$	0.868	0.868	0.868
H4 : HN	Stochasticity (no sea-ice) + $p(\text{future catastrophe}) = p^*$	0.649	0.649	0.649
H5 : HN	Stochasticity (no sea-ice) + no future catastrophes	0.707	0.707	0.707

Table 4.2 continued

Trial	Description	D10: 1+ Relative Increase		
		5%	5%	5%
H0 : BN	Deterministic + no future catastrophes	0.947	0.947	0.947
H1 : BN	Environmental stochasticity + no future catastrophes	0.973	0.973	0.973
H2 : BN	Environmental stochasticity + $p(\text{future catastrophe}) = 0.0625$	0.954	0.954	0.954
H3 : BN	Environmental stochasticity + $p(\text{future catastrophe}) = p^*$	0.960	0.960	0.960
H4 : BN	Stochasticity (no sea-ice) + $P(\text{future catastrophe}) = p^*$	0.807	0.807	0.807
H5 : BN	Stochasticity (no sea-ice) + no future catastrophes	0.846	0.846	0.846
H0 : HN	Deterministic + no future catastrophes	0.913	0.913	0.913
H1 : HN	Environmental stochasticity + no future catastrophes	0.951	0.951	0.951
H2 : HN	Environmental stochasticity + $p(\text{future catastrophe}) = 0.0625$	0.927	0.927	0.927
H3 : HN	Environmental stochasticity + $p(\text{future catastrophe}) = p^*$	0.932	0.932	0.932
H4 : HN	Stochasticity (no sea-ice) + $p(\text{future catastrophe}) = p^*$	0.725	0.725	0.725
H5 : HN	Stochasticity (no sea-ice) + no future catastrophes	0.776	0.776	0.776

Table 4.2 continued

Trial	Description	N9: Avg. Need Satisfaction		
		5%	5%	5%
H0 : BN	Deterministic + no future catastrophes	1.000	1.000	1.000
H1 : BN	Environmental stochasticity + no future catastrophes	1.000	1.000	1.000
H2 : BN	Environmental stochasticity + $p(\text{future catastrophe}) = 0.0625$	1.000	1.000	1.000
H3 : BN	Environmental stochasticity + $p(\text{future catastrophe}) = p^*$	1.000	1.000	1.000
H4 : BN	Stochasticity (no sea-ice) + $P(\text{future catastrophe}) = p^*$	1.000	1.000	1.000
H5 : BN	Stochasticity (no sea-ice) + no future catastrophes	1.000	1.000	1.000
H0 : HN	Deterministic + no future catastrophes	0.971	0.971	0.971
H1 : HN	Environmental stochasticity + no future catastrophes	0.974	0.974	0.974
H2 : HN	Environmental stochasticity + $p(\text{future catastrophe}) = 0.0625$	0.973	0.973	0.973
H3 : HN	Environmental stochasticity + $p(\text{future catastrophe}) = p^*$	0.973	0.973	0.973
H4 : HN	Stochasticity (no sea-ice) + $p(\text{future catastrophe}) = p^*$	0.959	0.959	0.959
H5 : HN	Stochasticity (no sea-ice) + no future catastrophes	0.964	0.964	0.964

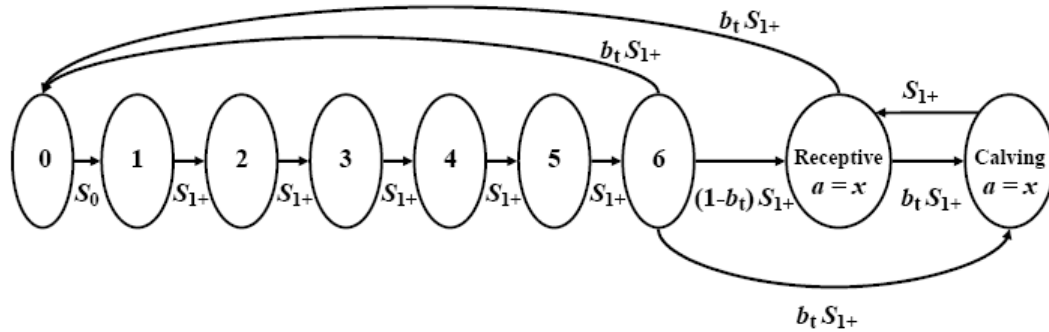


Figure 4.1: Life cycle graph of the model used to track the number of females in each reproductive stage through time. This life cycle refers to the underlying deterministic model, with transition probabilities shown as functions of life history parameters. However, it should be noted that the survival and birth rates were modified to be stochastic in the all analyses except for ‘H0’. The arrow from immature to calf arises because some juveniles may mature and give birth (i.e. become pregnant at first estrous) during the projection interval from time t to $t+1$.

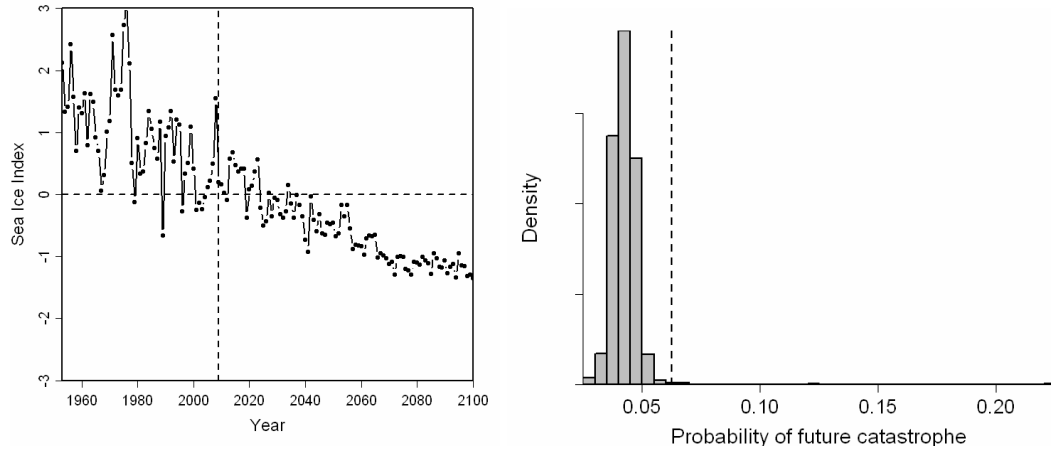


Figure 4.2: (Left panel) The standardized index for the March-April average sea-ice area covering the Bering Sea. The vertical dashed line denotes 2009 and the start of that portion of the time series which is based on the ensemble global climate model mean predictions provided by Overland and Wang (2007). Prior to 2009, the time series is based on the HadSST observations of sea-ice (Rayner *et al.*, 2003). The horizontal dashed line at zero is shown for reference; positive values indicate years with greater than average sea-ice over the entire time period and vice-versa. (Right panel) The distribution for the probability of future catastrophe. This distribution is conditioned on the number of years for which the depletion of each trajectory is greater 0.40 during 1930-2008, divided by 2 (the number of years with observed catastrophes, corresponding to 1999 and 2000). The dashed vertical line denotes the probability as calculated from the strandings index (Brownell *et al.*, 2007)

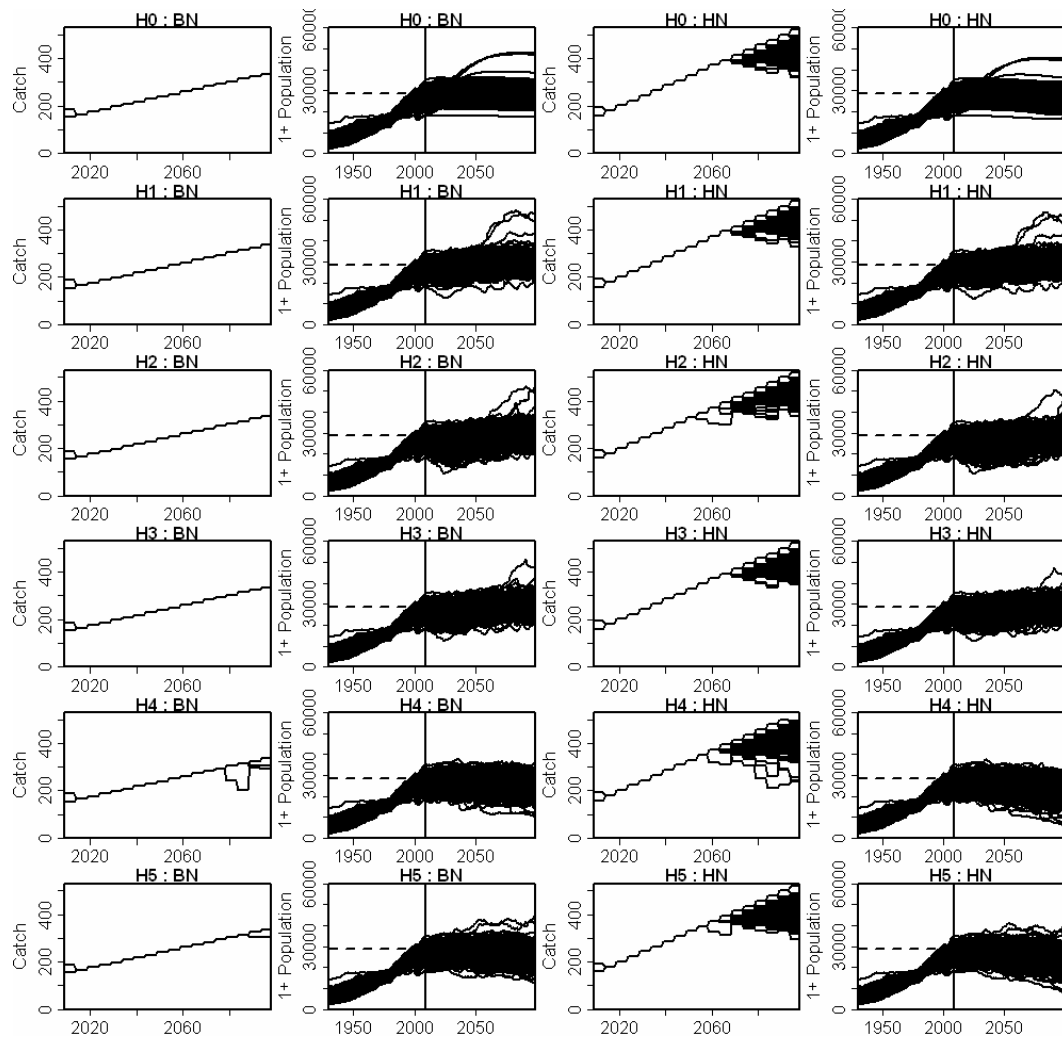


Figure 4.3: Time-trajectories of future catches (first and third columns) and population trajectories from 1930-2098 (second and fourth columns) for the twelve scenarios (Table 4.1). The left and right two columns are respectively for a final need levels of 340 and 530 whales per year. The results for each simulation are plotted as an individual line (e.g., a single visible line for catches represents a series of years where future catches were identical across scenarios).

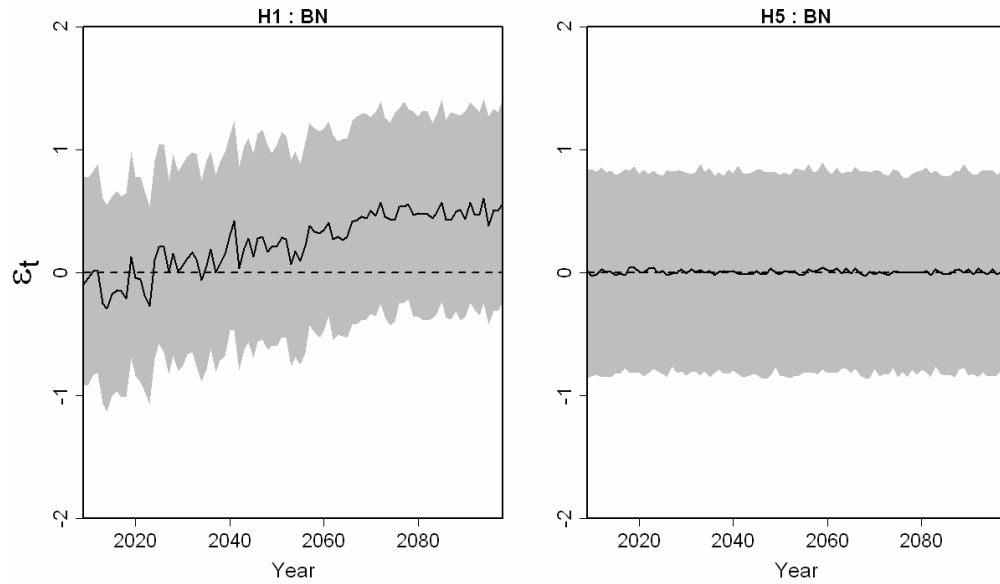


Figure 4.4: The time-trajectories of future process error residuals ε_t for a case where these residuals are a function of future sea-ice (H1:BN; left panel) and where they are independent of the sea-ice index (H5: BN; right panel). The annual median is plotted as the solid line, the 5th and 95th percentiles are shaded in gray and the horizontal dashed line at zero is shown for reference.

CONCLUSIONS

The preceding chapters quantify several key aspects of uncertainty in the context of marine mammal stock assessments. A Bayesian estimation framework was used for each example, which allowed the various dimensions of uncertainty to be integrated in the analyses and provided a logical foundation for calculating the probability that management strategies can achieve their desired objectives. Several case studies were examined, and it was found that existing management strategies are likely to be quite robust to the uncertainties considered when relatively a long time series of abundance estimates is available (i.e., BCB bowheads and ENP gray whales). On the other hand, the results of risk assessments were sensitive to how uncertainty was treated when constructing a coherent joint prior distribution that respects biological realism if abundance data are scarce (EG walrus).

A population dynamics model represents a working hypothesis with respect to the mechanisms that lead to changes in abundance. In common with all hypotheses, these models are based on underlying assumptions about the processes that drive the dynamics of the population. Therefore, model uncertainty can be thought of as uncertainty in which processes are most important in determining the population dynamics. Model selection methods are analogous to hypothesis testing; certain models may be rejected if they are not consistent with observations. However, given the relatively low growth rates of marine mammal populations and the inherent difficulties in estimating abundance for most populations of marine mammals, it is likely that a range of plausible models and assumptions will be consistent with the observed data. That is, available data may not provide much power to reject alternative hypotheses about the population dynamics. In Chapter 1, the assessment of the BCB stock of bowhead whales represented a situation for which alternative models were able to fit the data nearly equally well, but resulted in quite different estimates of stock status and sustainable catch levels. Each model

represented a variation on an underlying set of assumptions with respect to the historical catch record, carrying capacity and density dependence. Given available data, there was no evidence for rejecting the different assumptions underlying each model. Therefore, Bayesian model averaging was used to take this uncertainty into account, while also integrating parameter uncertainty and providing a weighted average of the model results based on their respective abilities to fit the data.

By using Bayesian model averaging, the assessment results presented in this case were able to incorporate uncertainty regarding assumptions such as the historical catch history is known without error and that the carrying capacity of the environment has remained unchanged since the middle of the 19th century for BCB bowheads. While taking into account the uncertainty in these assumptions is likely an improvement over previous approaches for this assessment which ignored such uncertainty, care must be given when considering the set of alternative models which are considered. If the candidate models are not biologically plausible or otherwise inappropriate, then the resulting estimates of management quantities may also be misleading, even if model uncertainty is taken into account. Munch *et al.* (2005) address this concern and provide a promising development in the treatment of model uncertainty based on Bayesian nonparametric modeling and using fishery stock-recruitment models as an example. This approach is similar to Bayesian model averaging in that alternative models are weighted by their ability to fit available data, but provides a framework which integrates model uncertainty over a continuous range of models spanning the biologically plausible stock-recruitment model space. This approach could potentially be incorporated into stock assessment frameworks to take into account model uncertainty (e.g., uncertainty in the assumed form of density dependence) while also accounting for parameter and observation uncertainty. Further, as fisheries management continues to move towards an ecosystem approach based on inherently complex models, the need to

account for model uncertainty is likely to become increasingly important (Hill *et al.*, 2007).

Bayesian methods have become progressively more popular in fishery science because they provide a natural framework for integrating and quantifying multiple sources of uncertainty. As discussed above, model uncertainty is one potentially important level in this hierarchy. Another fundamental level of uncertainty is expressed by the joint prior distribution for model parameters, which represents the uncertainty in those quantities prior to observing available data. Chapter 2 explored a subtle, but potentially important, aspect of constructing a coherent joint prior distribution, given an imposed constraint which ensured biological realism for the functionally related parameters in age-structured population dynamics models. It is evident from the results of this research that uncertainties resulting from alternative approaches for constructing the joint prior are unlikely to affect the results of assessments when informative time series of abundance estimates are available. In contrast, the results of a risk assessment for a data-poor stock were shown to be sensitive to the alternative approaches for creating the joint prior distribution. Unfortunately, there are no objective criteria on which to base a recommendation for adopting one approach to constructing a coherent joint prior over another. However, the results of Chapter 2 serve as a reminder that, when information on the trend in abundance is lacking (as is the case for many stocks of marine mammals), age-structured models are not necessarily the best approach for providing management advice. Instead, alternative methods should be considered for calculating sustainable limits of human-caused mortality. For example, the ‘Potential Biological Removal’ (PBR) rule adopted by the U.S. National Marine Fisheries Service (Wade, 1998) may be a better candidate than data-intensive age-structured models for determining precautionary catch limits for EG walrus. Indeed, the lack of sufficient data to inform more complex models for many stocks of marine mammals was exactly the motivation that led to the development of the

PBR as a simple, yet robust, method for assessing the appropriateness of removal levels (Taylor *et al.*, 2000).

Constructing coherent joint prior distributions which respect biological realism is a challenge that is not unique to marine mammal stock assessments. For example, Punt and Hilborn (1997) provide an example of how an incoherent distribution results from placing priors on both unexploited biomass and depletion (the ratio of current to unexploited biomass) in a biomass dynamics model for the western stock of New Zealand hoki (*Macruronus novaezelandiae*). Indeed, specifying contradictory priors is a potential pitfall during the development of any Bayesian analysis. While care must be taken in constructing a coherent joint prior, the real challenge is more likely to involve the necessary imposition of constraints on parameter space to ensure biologically realistic parameter values. This is a challenge which extends beyond marine mammal stock assessments and even the broader realm of fisheries models. It is a consideration for any application in the general arena of natural resource modeling whenever biological constraints on parameter values form part of the prior knowledge about the system in question.

Although walrus and many other marine mammals are extremely difficult to survey because they are distributed widely over large areas in relatively inaccessible habitats, some populations have been successfully monitored for long time periods e.g. the ENP gray whale. In Chapter 3, a modeling framework was developed for this stock, which allowed for the incorporation of a hypothesized relationship between sea-ice and population dynamics while also taking into account the mortality event in 1999 and 2000. This research represented the first time that the extent of this mortality event was quantified and integrated into a stock assessment. In this case, given the uncertainty in the magnitude of the mortality event and the availability of relevant data on fluctuations in birth and death rates (i.e., calf production and strandings data), the stochastic modeling framework developed in Chapter 3 is an appropriate and necessary approach for

providing estimates of stock status, because failure to take the mortality event into account is likely to lead to overly optimistic estimates of stock status. Therefore, the framework developed in Chapter 3 should be considered in addition to standard deterministic models during future assessments of this population.

The role of environmental variability on population dynamics is an area of special concern for management of marine mammals which are found in the Arctic, because forecasted climate change and anticipated changes in sea-ice are expected to be dramatic in this region. In Chapter 4, the modeling framework developed in Chapter 3 was used as an operating model in order to test the Gray Whale *SLA* of the IWC, given future predictions of annual sea-ice area on the Bering Sea feeding grounds. The estimated relationship between sea-ice and calf production was used to extrapolate future population dynamics based on the forecasted decrease in sea-ice. Future catches were determined by the *SLA* under different levels of future aboriginal subsistence need and account was taken of observation error in future estimates of abundance as generated using the operating model. While the *SLA* performed very well under the scenarios considered in these analyses, there is still much uncertainty about how changes in sea-ice (or other environmental conditions) will affect future population dynamics. At present, the level of information about the affects of environmental variability on population dynamics is largely correlative in nature. That is, the underlying mechanisms responsible for fluctuations in birth and survival rates are not well understood. Although a plausible explanation has been hypothesized (i.e., that sea-ice may act as a physical barrier to prime feeding habitat), it is not straightforward to predict how other likely changes in the environment resulting from reductions in sea-ice will interact with the mechanisms that are currently in force. Therefore, while the results of Chapter 4 indicate that the Gray Whale *SLA* is robust to predicted changes in arctic sea-ice, this conclusion must be tempered by uncertainty in the underlying assumption that current ecological processes will remain unchanged in the future,

especially when so many other fundamental changes in ecosystems are expected as a result of climate change.

As continuing research provides more insight into the mechanisms underlying the impacts of environmental variability on population dynamics of ENP gray whales, the modeling framework developed here will provide a good basis for integrating such information into assessments and management strategy evaluations. One of the appealing attributes of this framework for incorporating environmental data is its flexibility. For example, it is simple to substitute alternative environmental data during the model fitting process. Likewise, this framework can be applied to other stocks for which environmental fluctuations are hypothesized to be an important determinant of variability in population dynamics or behavior, and could prove to be useful in modeling changes in the availability of animals with respect to surveys or fishery catches. Additionally, it would be relatively straightforward to take into account hypothesized changes in the interaction between environmental variability and population dynamics in future projections (e.g., a loss of sea-ice might be beneficial for the ENP gray whales until the point at which the loss of sea-ice might result in negative effects owing to other impacts on the ecosystem).

Such considerations bring the research presented here full-circle. That is, the underlying uncertainty in the assumption that a certain index of environmental conditions is related to population dynamics and that if such a relationship exists, that current processes will remain unchanged in the future, are really manifestations of model uncertainty. Further, instead of assuming that current relationships will persist independently of other ecological changes, it may be preferable to enforce constraints on the degree to which future environmental conditions may affect population dynamics to respect biological realism. Explicitly accounting for this last point is beyond the scope of this dissertation, but does illustrate how the

various issues dealt with in this research are potentially inter-related in the broader context of providing management advice in the face of scientific uncertainty.

In summary, this dissertation explores several notable categories of uncertainty in marine mammal stock assessments. This research serves to improve the current understanding of population dynamics by incorporating and quantifying these various sources of uncertainty, and hence it also aims to ensure that resulting management advice is robust to these issues. In addition, the methods developed for incorporating environmental variability as well as the issues of model uncertainty and constructing coherent joint prior distributions which respect biological realism, are matters likely relevant to a large range of ecological modeling efforts. Therefore, the methods applied here are broadly applicable, and also serve as a basis for providing the best available management advice for marine mammals as well as other exploited natural resources.

REFERENCES

- ACIA [Arctic Climate Impact Assessment]. 2004. Impacts of a warming Arctic: Arctic Climate Impact Assessment. Cambridge University Press, Cambridge, UK.
- A'mar, Z.T., Punt, A.E. and Dorn, M.W. 2009. The evaluation of two management strategies for the Gulf of Alaska walleye Pollock under climate change. *ICES J.Mar.Sci.* 66: 1614-32.
- Allen, K.R. 1976. A more flexible model for baleen whale populations. *Rep. int. Whal. Commn.* 26: 247-63.
- Alvarez-Flores, C.M. 2002. Uncertainty in the management of activities affecting marine mammal populations: the tuna-dolphin conflict, a case study. Ph.D. thesis, University of Washington, Seattle, WA. 166 pp.
- 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.
- Barlow, J., and Boveng, P. 1991. Modeling age-specific mortality patterns for marine mammal populations. *Mar. Mammal Sci.* 7: 50-65.
- Best, P.B. 1993. Increase rates in severely depleted stocks of baleen whales. *ICES J. Mar. Sci.* 50: 169-86.
- Bockstoce, J.R. and Botkin, D.B. 1983. The historical status and reduction of the western Arctic bowhead whale (*Balaena mysticetus*) population by the pelagic whaling industry, 1848-1914. *Rep. int. Whal. Commn.* (Special Issue 5):107-41.
- Born, E.W. 2001. Reproduction in female Atlantic walruses (*Odobenus rosmarus rosmarus*) from north-western Greenland. *J. Zool.* 255: 165-74.
- Born, E.W., Dietz, R., Heide-Jørgensen, M., and Knutsen, L. 1997. Historical and present status of the Atlantic walrus (*Odobenus rosmarus rosmarus*) in eastern Greenland. *Meddelelser om Grønland. Bioscience.* 46: 1-73.
- Braham, H. W. 1989. Eskimos, Yankees, and bowheads. *Oceanus* 32(1): 54-62.

- Braham, H. W. 1995. Sex and size composition of bowhead whales landed by Alaskan Eskimo whalers. Hunting the largest animals: native whaling in the western Arctic and subarctic. A. P. McCartney. Edmonton, Alberta, Canada, Canadian Circumpolar Institute, University of Alberta: 281-313.
- Brandon, J. R., and Wade, P. R. 2006. Assessment of the Bering-Chukchi-Beaufort Seas stock of bowhead whales using Bayesian model averaging. *J. Cetacean Res and Manage.* 8: 225-39.
- Bravington, M.V. 1996. An appraisal of Bayesian synthesis, with suggested modifications and diagnostics. *Rep. Int. Whaling Comm.* 46: 531-40.
- Breen, P.A., Hilborn, R., Maunder, M.N., and Kim, S.W. 2003. Effects of alternative control rules on the conflict between a fishery and a threatened sea lion (*Phocartos hookeri*). *Can. J. Fish. Aquat. Sci.* 60: 527-41.
- Breen, P.A., and Kim, S.W. 2006. An integrated Bayesian evaluation of Hooker's sea lion bycatch limits. In *Sea Lions of the World*, pp. 471-494. Ed. by A. W. Trites, S. K. Atkinson, D. P. DeMaster, L. W. Fritz, T. S. Gelatt, L. D. Rea and K. M. Wynne. Alaska Sea Grant College Program, University of Alaska Fairbanks, Fairbanks.
- Breiwick, J. M. and Braham, H. W. 1990. Historical population estimates of bowhead whales: sensitivity to current population size. *Rep. Int. Whaling Comm.* 40: 423-26.
- Breiwick, J.M., Eberhardt, L.L. and Braham, H.W. 1984. Population dynamics of western Arctic bowhead whales (*Balaena mysticetus*). *Can. J. Fish. Aquat. Sci.* 41:484-496.
- Brownell Jr., R.L., Makeyev, C.A.F and Rowles, T.K. 2007. Stranding trends for eastern gray whales, *Eschrichtius robustus*: 1975-2006. SC/59/BRG40 presented to the IWC Scientific Committee, May 2007, Anchorage, U.S.A. (unpublished).
- Buckland, S.T., Burnham, K.P. and Augustin, N.H. 1997. Model selection: an integral part of inference. *Biometrics* 53:603-18.
- Buckland, S.T., Newman, K.B., Thomas, L., and Koesters, N.B. 2004. State-space models for the dynamics of wild animal populations. *Ecol. Model.* 171: 157-75.

- Burnham, K.P., and Anderson, D.R. 2002. Model Selection and Inference: A Practical Information-theoretic Approach. 2nd Ed. Springer-Verlag, New York.
- Butterworth, D.S. 1995. On reservations concerning the results of the Bayesian synthesis analysis, given the approach used to develop prior distributions for natural mortality rates. *Rep. int. Whal. Commn.* 45: 163-164.
- Butterworth, D.S. 2007. Why a management procedure? Some positives and negatives. *ICES J.Mar.Sci.* 64: 613-17.
- 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., David, J.H.M, McQuaid, L.H., and Xulu, S.S. 1987. Modeling the population dynamics of the South African fur seal (*Arctocephalus pusillus pusillus*). In Status, biology, and ecology of fur seals. Ed. by J. P. Croxall and R. L. Gentry. Proceedings of an International Symposium and Workshop, Cambridge, England, 23-27 April 1984. NOAA Technical Rep. NMFS 51: 141-164, U.S. Dept. of Commerce.
- Butterworth, D.S., De Oliveira, J.A.A., & Cochrane, K.L. 1993. Current initiatives in refining the management procedure for the South African anchovy resource. p.439-473, in: G. Kruse, D.M. Eggers, R.J. Marasco, C. Pautzke and T.J. Quinn II (eds). Proceedings of the International Symposium on Management Strategies for Exploited Fish Populations. Alaska Sea Grant College Program Report No. 93-02. University of Alaska, Fairbanks.
- Caughley, G. 1966. Mortality patterns in mammals. *Ecology*. 47: 906-918.
- Chivers, S. 1999. Biological indices for monitoring population status of walrus evaluated with an individual-based model. In Marine Mammal Survey and Assessment Methods, pp. 239-247. Ed. by G. W. Garner, S. C. Amstrup, J. L. Laake, B. F. L. Manly, L. L. McDonald, and D. G. Robertson. Proceedings of the Symposium on Surveys, Status & Trends of Marine Mammal Populations. Seattle, Washington, USA, 25-27 February 1998. Balkema, A.A., Rotterdam, Brookfield, The Netherlands.
- Clyde, M. and George, E.I. 2004. Model uncertainty. *Stat. Sci.* 19: 81-94.
- Cooke, J.G, Weller, D.W., Bradford, A.L., Burdin, A.M. and Brownell Jr., R.L. 2007. Population assessment of western gray whales in 2007. SC/59/41

presented to the IWC Scientific Committee, May 2007, Anchorage, U.S.A. 10pp. (unpublished).

- Da Silva, C.Q., J. Zeh, D. Madigan, J. Laake, D. Rugh, L. Baraff, W. Koski, and G. Miller. 2000. Capture-recapture estimation of bowhead whale population size using photo-identification data. *J. Cetacean Res and Manage.* 2(1):45-61.
- 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* 39:150-51.
- Dinesman, L.G. and Savinetsky, A.B. 2003. Secular dynamics of the prehistoric catch and population size of baleen whales off the Chukchi Peninsula, Siberia. pp. 137-166 *In: A.P. McCartney (ed.) Indigenous Ways to the Present: Native Whaling in the Western Arctic.* Canadian Circumpolar Institute Press, Edmonton.
- Durant J.M., Hjermann, D.Ø., Anker-Nilssen, T., Beaugrand, G., Mysterud, A., Pettorelli, N. and Stenseth, N.C. 2005. Timing and abundance as key mechanisms affecting trophic interactions in variable environments. *Ecol Lett* 8: 952-58.
- Durban, J.W., Elston, D.A., Ellifrit, D.K., Dickson, D., Hammond, P.S., and Thompson, P.M. 2005. Multisite mark-recapture for cetaceans: population estimates with Bayesian model averaging. *Mar. Mamm. Sci.* 21(1): 80-92
- Eberhardt, L.L. 1977. Optimal policies for conservation of large mammals with special reference to marine ecosystems. *Environ. Conserv.* 4: 205-12.
- Ellison, A.M. 1996. An introduction to Bayesian inference for ecological research and environmental decision-making. *Ecol. Appl.* 6: 1036-1046.
- Euler, L. 1770. A general investigation into the mortality and multiplication of the human species. *Theor. Popul. Biol.* 1: 307-14. (Originally published 1760).
- Fay, F. 1982. Ecology and biology of the Pacific walrus, *Odobenus rosmarus divergens* Illeger. North American Fauna U.S. Department of the Interior Fish and Wildlife Service 74: 1-279.
- Fay, G. 2004. A Bayesian stochastic metapopulation model for Steller sea lions in Alaska. M.Sc. thesis, University of Washington, Seattle, WA. 253 pp.

- Fay, G., and Punt, A.E. 2006. Modeling spatial dynamics of steller sea lion (*Eumetopias jubatus*) using maximum likelihood and Bayesian methods: evaluating causes for population decline. In *Sea Lions of the World*, pp. 405-433. Ed. by A. W. Trites, S. K. Atkinson, D. P. DeMaster, L. W. Fritz, T. S. Gelatt, L. D. Rea and K. M. Wynne. Alaska Sea Grant College Program, University of Alaska Fairbanks, Fairbanks.
- Forcada, J., Trathan, P.N., Reid, K. and Murphy, E.J. 2005. The effects of global climate variability in pup production of Antarctic fur seals. *Ecology*. 86: 2408-17.
- French, S. 1985. Group consensus probability distributions: a critical survey. In *Bayesian Statistics 2*, pp. 183-202. Ed. by J.M. Bernardo, M.H. DeGroot, D.V. Lindley and A.F.M. Smith. *Proceedings of the Second Valencia International Meeting*, 6-10 September 1983. Elsevier Science Publishers, Amsterdam.
- Gaillard, J.M., Festa-Bianchet, M., Yoccoz, N. G., Loison, A. and Toigo, C. 2000. Temporal variation in fitness components and population dynamics of large herbivores. *Annu. Rev. Ecol. Syst.* 31: 367-93.
- Gelman, A., Carlin, J.B., Stern, H.S. and Rubin, D.B. 2004. *Bayesian Data Analysis*. Chapman & Hall/CRC. 2nd Edn., Boca Raton, FL. xxv+668 pp.
- Genest, C., and Zidek, J.V. 1986. Combining probability distributions: a critique and an annotated bibliography. *Stat. Sci.* 1: 114-35.
- George, J. C., Philo, L. M., Carroll, G.M. and Albert, T.F. 1988. 1987 subsistence B-harvest of bowhead whales, *Balaena mysticetus*, by Alaskan Eskimos. *Rep. Int. Whal. Comm.* 38: 389-92.
- 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 acid racemization. *Can. J. Zool.* 77(4): 571-580.
- George, J. C., Zeh, J., Suydam, R. and Clark, C. 2004. Abundance and population trend (1978-2001) of western arctic bowhead whales surveyed near Barrow, Alaska. *Mar. Mamm. Sci.* 20(4): 755-73
- George, J.C., Nicolson, S., Drobot, S., Maslanik, J. and Rosa, C. 2009. Update on SC/57/E13: Progress report: sea ice density and bowhead whale body condition. SC/F09/CC1. Presented to the IWC Climate Change Workshop, Siena, Italy, Feb. 2009. [Available from the Secretariat]

- Geweke, J. 1992. Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In *Bayesian Statistics 4* (ed JM Bernardo, JO Berger, AP Dawid and AFM Smith). Clarendon Press, Oxford, UK.
- Givens, G.H. 1993. A Bayesian framework and importance sampling methods for synthesising multiple sources of evidence and uncertainty linked by a complex mechanistic model. Ph.D. thesis, University of Washington, Seattle, WA. 157 pp.
- Givens, G.H. 1999. Bayesian population dynamics modeling using uncertain historical catch, surveyed abundances, and biological priors. 1. In *Marine Mammal Survey and Assessment Methods*, pp. 225-238. Ed. by G. W. Garner, S. C. Amstrup, J. L. Laake, B. F. L. Manly, L. L. McDonald, and D. G. Robertson. Proceedings of the Symposium on Surveys, Status & Trends of Marine Mammal Populations. Seattle, Washington, USA, 25-27 February 1998. Balkema, A.A., Rotterdam, Brookfield, The Netherlands.
- Givens, G.H., and Roback, P.J. 1999. Logarithmic pooling of priors linked by a deterministic simulation model. *J. Comput. Graph. Stat.* 8: 452-78.
- Givens, G.H., and Thompson, S.E. 1996. Alternative Bayesian synthesis approaches to Bering-Chukchi-Beaufort Seas bowhead whale stock assessment: uncertainty in historic catch and hitting with fixed MSYR. *Rep. int. Whal. Commn.* 46: 509-30.
- Givens, G.H., Raftery, A.E. and Zeh, J.E. 1993. Benefits of a Bayesian approach for synthesizing multiple sources of evidence and uncertainty linked by a deterministic model. *Rep. int. Whal. Commn.* 43:495-500.
- Givens, G.H., Zeh, J.E. and Raftery, A.E. 1995. Assessment of the Bering-Chukchi-Beaufort Seas stock of bowhead whales using the BALEEN II model in a Bayesian synthesis framework. *Rep. int. Whal. Commn.* 45: 345-64.
- Goodman, D. 1984. Statistics of reproductive rate estimates, and their implications for population projection. *Rep. int. Whal. Commn.* (Special Issue 6): 161-73.
- Goebel, M.E. 2002. Northern fur seal lactation, attendance and reproductive success in two years of contrasting oceanography. Ph.D. thesis. University of California, Santa Cruz.
- Green, C.H., Pershing, A.J., Kenney, R.D. and Jossi, J.W. 2003. Impact of climate variability on the recovery of endangered North Atlantic right whales. *Oceanography* 16: 98-103.

- Guan, B.T. 2000. Effects of correlation among parameters on prediction quality of a process-based forest growth model. *Forest Sci.* 46: 269-76.
- Gulland, F.M.D., H. Pérez-Cortés M., J. Urbán R., L. Rojas-Bracho, G. Ylitalo, J. Weir, S.A. Norman, M.M. Muto, D.J. Rugh, C. Kreuder, and T. Rowles. 2005. Eastern North Pacific gray whale (*Eschrichtius robustus*) unusual mortality event, 1999-2000. In NOAA Technical Memorandum NMFS-AFSC-150.[www.afsc.noaa.gov/Publications/AFSC-TM/NOAA-TM-AFSC-150.pdf]
- Griewank, A. and Corliss, G.F. 1991. Automatic differentiation algorithms: theory, implementation, and application. SIAM, Philadelphia.
- Harley, S.J. and Maunder, M.N. 2004. Status of bigeye tuna in the eastern Pacific Ocean in 2002 and outlook for 2003. *Inter-Amer. Trop. Tuna Comm., Stock Assessment Report*, 4: 120-286.
- Hastings, W.K. 1970. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*. 57: 97-109.
- Heidelberger, P. and Welch, P.D. 1983. Simulation run length control in the presence of an initial transient. *Opns. Res.* 31: 1109-44.
- Hilborn, R., and Walters, C.J. 1992. Quantitative Fisheries Stock Assessment: Choice, Dynamics and Uncertainty. Chapman and Hall, London. 570 pp.
- Hill, S.L, Watters, G.M., Punt, A.E., McAllister, M.K., Le Quere, C. and Turner, J. 2007. Model uncertainty in the ecosystem approach to fisheries. *Fish Fish.* 8: 315-36.
- Hoeting, J.A., Madigan D., Raftery A.E. and Volinsky, C.T. 1999. Bayesian Model Averaging: A Tutorial. *Statistical Sciences* 14 (4):382-417.
- Hoyle, S.D., and Maunder, M.N. 2004. A Bayesian integrated population dynamics model to analyze data from protected species. *Anim. Biodivers. Conserv.* 27: 247-66.
- Iman, R.L., and Conover, W.J. 1982. A distribution-free approach to inducing rank correlation among input variables. *Commun. Stat: B-Simulat.* 11: 311-34.
- IPCC [Intergovernmental Panel on Climate Change]. 2007. Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M.

Tignor and H.L. Miller (eds.). Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 996 pp.

International Whaling Commission (IWC). 1993. Report of the Special Meeting of the Scientific Committee on the Assessment of Gray Whales, Seattle, 23-27 April 1990. *Rep. int. Whal. Commn.* 43:241-59.

International Whaling Commission (IWC). 1994. Chairman's Report of the Forty-Fifth Annual Meeting, Appendix 12. Resolution on research on the environment and whale stocks. *Rep. int. Whal. Commn.* 44:35.

International Whaling Commission (IWC). 1995. Report of the Scientific Committee, Annex F, Report of the Sub-Committee on Aboriginal Subsistence Whaling. *Rep. int. Whal. Commn.* 45: 142-64.

International Whaling Commission (IWC). 1997. Report of the IWC Workshop on Climate Change and Cetaceans. *Rep. int. Whal. Commn.* 47:293-313.

International Whaling Commission (IWC). 1998. Report of the Scientific Committee. Annex J. Report of the Sub-Committee on Aboriginal Subsistence Whaling. *Rep. int. Whal. Commn.* 48:237-48.

International Whaling Commission (IWC). 1999. Report of the Scientific Committee, Annex G. Report of the Sub-Committee on Aboriginal Subsistence Whaling. *J. Cetacean Res. Manage.* (Suppl.) 1:179-94.

International Whaling Commission (IWC). 2000. International Convention for the Regulation of Whaling, 1946. Schedule. Annual Report of the International Whaling Commission 1999, 77-90.

International Whaling Commission (IWC). 2003. Report of the Scientific Committee, Annex E. Report of the Standing Working Group on the Development of an Aboriginal Subsistence Whaling Management Procedure (AWMP). *J. Cetacean Res. Manage.* (Suppl.) 5: 154-255.

International Whaling Commission (IWC). 2003. Report of the Scientific Committee. *J. Cetacean Res. Manage.* (Suppl.) 5:1-92.

International Whaling Commission (IWC). 2003. Report of the Scientific Committee, Annex E. Report of the Standing Working Group on the Development of an Aboriginal Subsistence Whaling Management Procedure (AWMP). *J. Cetacean Res. Manage.* (Suppl.) 5: 154-255.

- International Whaling Commission (IWC). 2004. Chair's Report of the Fifty-Sixth Annual Meeting. Annual Report of the International Whaling Commission, *J. Cetacean Res. Manage.* (Suppl.) 4: 1–58.
- International Whaling Commission (IWC). 2005. Report of the Scientific Committee. Annex E. Report of the Standing Working Group (SWG) on the development of an Aboriginal Subsistence Whaling Management Procedure (AWMP). Appendix 3. Fishery Type 2: Implementation for Eastern North Pacific Gray Whales *J. Cetacean Res. Manage.* (Suppl.) 5:127-40
- International Whaling Commission (IWC). 2005. Report of the Scientific Committee, Annex F. Report of the Sub-Committee on Bowhead, Right and Gray Whales. *J. Cetacean Res. Manage.* (Suppl.) 7: 189-209.
- Jeffreys, H. 1961. *Theory of Probability*. 3rd Edn. Oxford University Press, London. i-vii+411pp.
- Kass, R.E. and Raftery, A.E. 1995. Bayes factors. *J. Am. Stat. Assoc.* 90: 773-95.
- Koski, W.R., Rugh, D.J., Punt, A.E., and Zeh, J. 2006. A new approach to estimating the length-frequency distribution of the Bering-Chukchi-Beaufort bowhead whale population using aerial photogrammetry data. *J. Cetacean Res. Manage.* 8: 45-54.
- Koski, W.R., George, J.C., and Zeh, J. 2008. A calf index for monitoring reproductive success in the Bering-Chukchi-Beaufort Seas bowhead whale (*Balaena mysticetus*) population. *J. Cetacean Res. Manage.* 10: 99-106.
- Leaper, R., Cooke, J., Trathan, P., Reid, K., Rowntree, V. and Payne, R. 2006. Global climate drives southern right whale (*Eubalaena australis*) population dynamics. *Biol. Lett.* 2: 289-92.
- Leslie, P. H. 1945. On the use of matrices in certain population mathematics. *Biometrika* 33: 183-212.
- Leslie, P. H. 1948. Some further notes on the use of matrices in population mathematics. *Biometrika* 35: 213-45.
- Lockyer, C. 1984. Review of baleen whale (Mysticeti) reproduction and implications for management. *Rep. int. Whal. Commn.* (Special Issue 6): 27-50.

- Lockyer, C. 1986. Body fat condition in north Atlantic fin whales, *Balaenoptera physalus*, and its relationship with reproduction and food resource. *Can. J. Fish. Aquat. Sci.* 43: 142-47.
- Lotka, A.J. 1907. Studies on the mode of growth of material aggregates. *Am. J. Sci.* 24: 199-216.
- Lowry, L.F. 2000. Marine mammal-sea ice relationships. Pages 123-30 in *Impacts of changes in sea ice and other environmental parameters in the Arctic: Report of the Marine Mammal Commission workshop*. Huntington. H.P. (ed.), Girdwood, Alaska: Marine Mammal Commission, Bethesda, MD.
- Mansfield, A. 1958. The biology of the Atlantic walrus, *Odobenus rosmarus rosmarus* (Linnaeus) in eastern Canada. Fisheries Research Board of Canada Manuscript Report Series (Biology). 653: 1-146.
- Maunder, M.N. and Watters, G.M. 2003. A general framework for integrating environmental time series into stock assessment models: model description, simulation testing and example. *Fish. Bull, US* 101: 89-99.
- Methot, R.D. 2000. Technical Description of the Stock Synthesis Assessment Program. NOAA Technical Memorandum, NMFS-NWFSC-43, 46pp. [http://www.nwfsc.noaa.gov/assets/25/5560_06162004_143158_tm43.pdf].
- Munch, S.B., Kottas, A. and Mangel, M. 2005. Bayesian nonparametric analysis of stock-recruitment relationships. *Can. J. Fish. Aquat. Sci.* 62: 1808-21.
- Nakamura, T., Ohnishi, S. and Matsumiya, Y. 1989. A Bayesian cohort model for catch-at-age data obtained from research takes of whales. *Rep. int. Whal. Commn.* 39: 375-82.
- Overland, J.E. and Wang, M. 2007. Future regional Arctic sea ice declines. *Geophys. Res. Lett.* 34: L17705.
- Patterson, K.R. 1999. Evaluating uncertainty in B-C-B control law catches using Bayesian Markov chain Monte Carlo virtual population analysis with adaptive rejection sampling and including structural uncertainty. *Can. J. Fish. Aquat. Sci.* 51: 208-221.
- Pella, J.J., and Tomlinson, P.K. 1969. A generalized stock production model. Inter-American Tropical Tuna Commission Bulletin, 13, No. 3, pp. 419-96.

- Perryman, W.L. and Lynn, M.S. 2002. Evaluation of nutritive condition and reproductive status of migrating gray whales (*Eschrichtius robustus*) based on analysis of photogrammetric data. *J. Cetacean Res. Manage.* 4: 155-64.
- Perryman, W.L., Donahue, M.A., Perkins, P.C. and Reilly, S.B. 2002. Gray whale calf production 1994-2000: are observed fluctuations related to changes in seasonal ice cover? *Mar. Mammal Sci.* 18: 121-44.
- Poole, D. and Givens, G. H. 2001. An exploratory assessment of the Bering-Chukchi-Beaufort Seas stock of bowhead whales using a stochastic population dynamics model. *J. Cetacean Res. Manage.* 3: 1-6.
- Poole, D., Givens, G. H. and Raftery, A.E. 1999. A proposed stock assessment method and its application to bowhead whales, *Balaena mysticetus*. *Fish. Bull.* 97(1): 144-52.
- Poole, D. and Raftery, A.E. 1998. Assessment of the Bering-Chukchi-Beaufort Seas stock of bowhead whales using Bayesian and full pooling Bayesian synthesis. Paper SC/50/AS6 presented to the IWC Scientific Committee, April 1998 (unpublished) 34pp.
- Poole, D., and Raftery, A.E. 2000. Inference for deterministic simulation models: the Bayesian melding approach. *J. Am. Stat. Assoc.* 95: 1244-55.
- Poole, M.M. 1984. Preliminary assessment of annual calf production in the gray whale, *Eschrichtius robustus*, from Pt Piedras Blancas, California. *Rep. int. Whal. Commn.* (Special Issue 6): 223-31.
- Press, S.J. 1989. Bayesian statistics: principles, models, and applications. John Wiley, N.Y. 237 pp.
- Press, W.H., Tevkolsky, S.A., Vetterling, W.T. and Flannery, B.P. 1992. Numerical Recipes in Fortran 2nd Ed. (Cambridge Univ. Press).
- Punt, A.E. 1999a. 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. 1999b. Results for two selected assessments of the Bering-Chukchi-Beaufort Seas stock of bowhead whales. Appendix 4 to Report of the Scientific Committee, Annex G. Report of the Sub-Committee on Aboriginal Subsistence Whaling. *J. Cetacean Res. Manage.* (Suppl.) 1:192-4.

- Punt, A.E. 2006. Assessing the Bering-Chukchi-Beaufort Seas stock of bowhead whales using abundance data together with data on length or age. *J. Cetacean Res. Manage.* 8: 127-37.
- Punt, A.E. 2009. A note on the modelling of MSY-related parameters when population dynamics are stochastic. *J. Cetacean Res. Manage.* 10: 183-89.
- Punt, A.E. and Breiwick, J.M. 2008. On standalone versions of the Bowhead and Gray Whale *SLAs*. Paper SC/M08/AWMP1 presented to the Workshop on Developing Assessment Methods and a Management Procedure for Greenlandic Fisheries, 26-29 March 2008, Copenhagen (unpublished). 11pp.
- Punt, A.E., and Butterworth, D.S. 1996. Further remarks on the Bayesian approach for assessing the Bering-Chukchi-Beaufort Seas stock of bowhead whales. *Rep. int. Whal. Commn.* 46: 481-91.
- Punt, A.E. and Butterworth, D.S. 1999. On assessment of the Bering-Chukchi-Beaufort Seas stock of bowhead whales (*Balaena mysticetus*) using a Bayesian approach. *J. Cetacean Res. Manage.* 1:53-71.
- Punt, A. E. and Butterworth, D.S. 2000. Why do Bayesian and maximum likelihood assessments of the Bering-Chukchi-Beaufort Seas stock of bowhead whales differ? *J. Cetacean Res. Manage.* 2: 125-33.
- Punt, A.E. and Butterworth, D.S. 2002. An examination of certain of the assumptions made in the Bayesian approach used to assess the eastern Pacific stock of gray whales (*Eschrichtius robustus*). *J. Cetacean Res. Manage.* 4: 99-110.
- Punt, A.E. and Donovan, G.P. 2007. Developing management procedures that are robust to uncertainty: lessons from the International Whaling Commission. *ICES J. Mar. Sci.* 64: 603-12.
- Punt, A.E. and Hilborn, R. 1997. Fisheries stock assessment and decision analysis: A review of the Bayesian approach. *Rev. Fish. Biol. Fish.* 7: 35-63.
- Punt, A.E., Allison, C. and Fay, G. 2004. An examination of assessment models for the Eastern North Pacific Gray whale based on inertial dynamics. *J. Cetacean Res. Manage.* 6:121-132.
- Radtke, P.J., Burk, T.E., and Bolstad, P.V. 2002. Bayesian melding of a forest ecosystem model with correlated inputs. *Forest Sci.* 48: 701-11.

- Raftery, A.E. and Lewis, S.M. 1992. One long run with diagnostics: Implementation strategies for Markov chain Monte Carlo. *Stat. Sci.* 7: 493-97.
- Raftery, A.E. and Poole, D. 1997. Bayesian synthesis assessment methodology for bowhead whales. Paper SC/49/AS2 presented to the IWC Scientific Committee (unpublished).
- Raftery, A.E. and Zeh, J.E. 1991. Bayes empirical Bayes estimation of bowhead whale population size based on the visual and acoustic census near Barrow, Alaska, in 1986 and 1988. Paper SC/43/PS8 presented to the IWC Scientific Committee, May 1991 (unpublished) 51pp.
- Raftery, A. E. and Zeh, J. E. 1993. Estimation of Bowhead Whale, *Balaena mysticetus*, Population Size. Case Studies in Bayesian Statistics. C. Gatsonis, J. S. Hodges, R. E. Kass and N. D. Singpurwalla. New York, Springer-Verlag:164-240.
- Raftery, A. E. and Zeh, J. E. 1998. Estimating bowhead whale population size and rate of increase from the 1993 census. *J. Am. Stat. Assoc.* 93: 451-63.
- Raftery, A.E., Givens, G.H., and Zeh, J.E. 1995. Inference from a deterministic population dynamics model for bowhead whales. *J. Am. Stat. Assoc.* 90: 402-16.
- Raftery, A.E., Gneiting, T. Balabdaoui, F. and Polakowski, M. 2005. Using Bayesian model averaging to calibrate forecast ensembles. *Mon. Weather Rev.* 133(5): 1155-74.
- Rayner, N.A., Parker, D.E., Horton, E. B., Folland, C. K., Alexander, L. V., Rowell, D. P., Kent, E. C. and Kaplan, A. 2003. Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res.* 108, No. D14, 4407.
- Reilly, S. B. and Barlow, J. 1986. Rates of increase in dolphin population size. *Fish. Bull.* 84: 527-33.
- Rice, D.W. and Wolman, A.A. 1971. The Life History and Ecology of the Gray Whale (*Eschrichtius robustus*). Stillwater, Oklahoma: American Society of Mammalogists, Special Publication No. 3.
- Rubin, D.B. 1988. Using the SIR algorithm to simulate posterior distributions. pp. 395-402 *In*: J.M. Bernardo, M.H. DeGroot, D.V. Lindley and A.F.M. Smith

- (eds.) *Bayesian Statistics 3: Proceedings of the Third Valencia International Meeting, June 1-5, 1987*. Clarendon Press, Oxford. 805pp.
- Rugh, D.J., Hobbs, R.C., Lerczak, J.A. and Breiwick, J.M. 2005. Estimates of abundance of the eastern North Pacific stock of gray whales (*Eschrichtius robustus*) 1997-2002. *J. Cetacean Res. Manage.* 7:1-12.
- Rugh, D., Breiwick, J., Hobbs, R., Shelden, K., and Muto, M. 2008. Eastern North Pacific gray whale abundance in the winter 2006-2007. SC/60/BRG6. Presented to the IWC Scientific Committee, June 2005. Santiago, Chile. (unpublished).
- Schirripa, M.J. and Colbert, J.J. 2006. Interannual changes in sablefish (*Anoplopoma fimbria*) recruitment in relation to oceanographic conditions within the California Current System. *Fish. Ocean.* 15: 25-36.
- Schirripa, M.J., Goodyear, P.C. and Methot, R.M. 2009. Testing different methods of incorporating climate data into the assessment of US West Coast sablefish. *ICES J. Mar. Sci.* 66: 1605-13.
- Schweder, T., and Hjort, N.L. 1996. Bayesian synthesis or likelihood synthesis - what does Borel's paradox say? *Rep. int. Whal. Commn.* 46: 475-79.
- Schweder, T. and Ianelli, J.N. 2000. Assessing the Bering-Chukchi-Beaufort Seas stock of bowhead whales from survey data, age-readings and photo-identifications using frequentist methods. Paper SC/52/AS13 presented to the IWC Scientific Committee, June 2000 (unpublished) 12pp.
- Simmonds, M.P. and Isaac, S.J. 2007. The impacts of climate change on marine mammals: early signs of significant problems. *Oryx* 41: 19-26.
- Smith, A.F.M. and Gelfand, A.E. 1992. Bayesian statistics without tears: a sampling-resampling perspective. *Am. Stat.* 46: 84-88.
- Smith, B.J. 2007. Bayesian Output Analysis Program (BOA) for MCMC v.1.1.5 [<http://www.public-health.uiowa.edu/boa>].
- Stirling, I., Lunn, N. and Iacozza, J. 1999. Long-term trends in the population ecology of polar bears in western Hudson Bay in relation to climatic change. *Arctic* 52: 294-306.
- Suydam, R., and George, C. 2004. Subsistence B-C-B of bowhead whales, *Balaena mysticetus*, by Alaska Eskimos, 1974 to 2003. Paper SC/56/BRG12 presented to the IWC Scientific Committee (unpublished).

- Taylor, B.L., Wade, P.R., De Master, D.P., and Barlow, J.. 2000. Incorporating uncertainty into management models for marine mammals. *Conserv. Biol.* 14: 1243-52.
- Thomas, L., Buckland, S.T., Newman, K.B., and Harwood, J. 2005. A unified framework for modelling wildlife population dynamics. *Australian and New Zealand Journal of Statistics*, 47: 19-34.
- Tynan, C.T. and DeMaster, D.P. 1997. Climatic change and marine mammals. *Arctic* 50:308-22.
- Wade, P.R. 1994. Abundance and population dynamics of two eastern Pacific dolphins, *Stenella attenuata* and *Stenella longirostris orientalis*. Ph.D. thesis, Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA. 255 pp.
- Wade, P.R. 1997. A revised gray whale analysis using both southbound total population counts and northbound calf counts. Paper SC/49/AS24 presented to the IWC Scientific Committee, September 1997, Bournemouth (unpublished). 23pp.
- Wade, P.R. 1998. Calculating limits to the allowable human-caused mortality of cetaceans and pinnipeds. *Mar. Mammal Sci.* 14: 1-37.
- Wade, P.R. 2002a. A Bayesian stock assessment of the eastern Pacific gray whale using abundance and harvest data from 1967 to 1996. *J. Cetacean Res. Manage.* 4:85-98.
- Wade, P.R. 2002b. Population Dynamics. Pages 974-979 in *Encyclopedia of Marine Mammals*, Perrin, W.F., Wursig, B., and Thewissen, J.G.M.(eds), Academic Press, San Diego.
- Wade, P.R. and Givens, G. 1997. Designing catch control laws that reflect the intent of aboriginal subsistence management principles. *Rep. int. Whal. Commn.* 47: 871-74.
- Wade, P.R. and Perryman, W. 2002. An assessment of the eastern gray whale population in 2002. Paper SC/54/BRG7 presented to the IWC Scientific Committee, April 2002, Shimonoseki, Japan (unpublished).
- Walsh, J.E. 1978. Data set on Northern Hemisphere sea ice extent, World Data Center A for Glaciology (Snow and Ice), Glaciological Data, Report GD-2, part 1, 49-51.

- Walters C. and Ludwig, D. 1994. Calculation of Bayes posterior probability distributions for key population parameters. *Can. J. Fish. Aquat. Sci.* 51: 713-22.
- Wang, S.J. 1998. Discussion of "Understanding Relationships Using Copulas," North American Actuarial Journal. 3: 137-141.
- Ward, E.J., Hilborn, R., Towell, R.G. and Gerber, L. 2007. A state-space mixture approach for estimating catastrophic events in time series data. *Can. J. Fish. Aquat. Sci.* 64: 899-910.
- Whitehead, H. 1997. Sea surface temperature and the abundance of sperm whale calves off the Galapagos Islands: Implications for the effects of global warming. *Rep. int. Whal. Commn.* 47: 941-44.
- Wilcox, C. and Eldred, B. 2003. The effect of density-dependent catastrophes on population persistence time. *J. Appl. Ecol.* 40: 859-71.
- Winship, A.J. and Trites, A.W. 2006. Risk of extirpation of Steller sea lions in the Gulf of Alaska and Aleutian Islands: a population viability analysis based on alternative hypotheses for why sea lions declined in western Alaska. *Mar. Mammal Sci.* 22: 124-55.
- Witting, L. 2003. Reconstructing the population dynamics of eastern Pacific gray whales over the past 150 to 400 years. *J. Cetacean Res. Manage.* 5: 45-54.
- Witting, L., and Born, E.W. 2005. An assessment of Greenland walrus populations. *ICES J. Mar. Sci.* 62: 266-84.
- Wolpert, R.L. 1995. Comment on 'Inference from a deterministic population dynamics model for bowhead whales'. *J. Am. Stat. Assoc.* 90: 426-27.
- Zeh, J.E. and Punt. A.E. 2005. Updated 1978-2001 abundance estimates and their correlations for the Bering-Chukchi-Beaufort Seas stock of bowhead whales. *J. Cetacean Res. Manage.* 7: 169-75.
- Zeh, J. E., George, J.C., Raftery, A.E. and Carroll, G.M. 1991. Rate of increase, 1978-1988, of bowhead whales, *Balaena mysticetus*, estimated from ice-based census data. *Mar. Mammal Sci.* 7(2): 105-22.
- Zeh, J., Poole, D., Miller, G., Koski, W., Baraff, L. and Rugh, D. 2002. Survival of bowhead whales, *Balaena mysticetus*, estimated from 1981-1998 photoidentification data. *Biometrics.* 58: 832-40.

Appendix A:
DERIVATION OF EQNS. 3.2 AND 3.3

In a given year, the realized stochastic birth rate can be written as a logistic function of a parameter related to the expected birth rate μ_t and process error ε_t that year:

$$b_t^* = [1 + \exp(-(\mu_t + \varepsilon_t))]^{-1} \quad \text{where:} \quad \varepsilon_t \sim N(0, \sigma_\varepsilon^2) \quad \text{A.1}$$

Further, let $\mu_t^* = \mu_t + \varepsilon_t$ where $\mu_t^* \sim N(\mu_t, \sigma_\varepsilon^2)$. Now, given a random variable generated from some underlying distribution $X \sim f(x)$ and a function of this random variable $Z = h(X)$, the expectation of the function can be written:

$$E[Z] = \int_{-\infty}^{\infty} h(x)f(x)dx \quad \text{A.2}$$

It is necessary to define the expectation of the stochastic birth rates as equal to the deterministic value from the Pella-Tomlinson model (Eqn. 3.1), i.e. $E[b_t^*] = b_t$ to model density dependence acting solely through the stochastic birth rate. Proceeding from the form of the expectation above, while noting that $b_t^* = h(\mu_t^*) = [1 + \exp(-\mu_t^*)]^{-1}$ and likewise, given its additive nature, that $\mu_t^* \sim N(\mu_t, \sigma_\varepsilon^2)$, substitute for $h(x)$ and $f(x)$ in equation A.2 and write the analytical expectation of the stochastic birth rates (Punt, 2008):

$$E[b_t^*] = \int_{-\infty}^{\infty} [1 + \exp(-\mu_t^*)]^{-1} \left[\frac{1}{\sigma_\varepsilon \sqrt{2\pi}} \exp\left(-\frac{(\mu_t^* - \mu_t)^2}{2\sigma_\varepsilon^2}\right) \right] d\mu_t^* \quad \text{A.3}$$

It was found through numerical methods that, the integral on the right side is well approximated by:

$$\approx \Phi\left(\frac{\mu_t}{\sqrt{2.76 + \sigma_\varepsilon^2}}\right) \quad \text{A.4}$$

where:

Φ is the standard normal cumulative distribution function; and
 2.76 is a value found through numerical minimization, which approximates this expectation.

Setting the expectation of the stochastic birth rates equal to the deterministic density dependent rate:

$$\Phi\left(\frac{\mu_t}{\sqrt{2.76 + \sigma_\varepsilon^2}}\right) = b_t \quad \text{A.5}$$

And finally, rearranging to solve this equation for μ_t and substituting back into equation A.1 yields Eqns. 3.2 and 3.3. Of course, for those years with added process error, the stochastic expectation will not necessarily be equal to the deterministic value.

VITA

John was born in Sacramento, California in 1976. He was raised in Davis with his younger brother, Patrick. His parents, John and Dorothy, introduced him to science through their respective involvements in the entomology and botany departments at UC Davis. He graduated from Davis High School in 1994. In 1998 he received his B.S. in Biology from UC San Diego, where he also captained the men's soccer team and spent the off-season in the ocean, surfing. He was hired during his senior year by Dr. Tim Gerrodette at the SWFSC of the NMFS, to assist in research of dolphin populations and the eastern tropical Pacific tuna fishery. During his 3 years as an employee of the SWFSC he became increasingly interested in the application of quantitative methods in conservation biology, and was also introduced to marine mammal surveys through field work in the eastern tropical Pacific and various locations along the west coast of the U.S. After traveling around the world by himself for a year, to surf and fly-fish in the southern hemisphere, he moved to Seattle for graduate school to study quantitative fisheries methods. Before enrolling at the UW, he worked as a marine mammal observer on aerial and ship-based surveys in the Arctic, and in the Aleutian Islands. Since that time, he has also been a marine mammal observer on surveys in the western Atlantic and the Yangtze River in China. Much of his dissertation research has been conducted for the International Whaling Commission, where John has been an Invited Participant during recent Scientific Committee meetings. He received his Ph.D. from the University of Washington, School of Aquatic and Fishery Sciences in 2009.