Review of contemporary cetacean stock assessment models

André E. Punt

School of Aquatic and Fishery Sciences, University of Washington, Seattle, WA 98195

ABSTRACT

Model-based methods of analysis are widely used to conduct cetacean stock assessments, and to provide the operating models on which management strategy evaluation is based. This paper reviews recent assessments and management strategy evaluations for cetacean populations, with a view towards establishing 'best practice' guidelines for such analyses. The models on which these analyses are based range from simple exponential trend models that ignore density-dependence to complex multi-stock age-sex- and stage-structured models that form the basis for management strategy evaluation. Most analyses assume that density-dependence is on calf survival (which implicitly includes maturity and pregnancy rate), but it could also impact the survival rate of adults or the age-at-maturity. Cetaceans seldom have more than one calf per female each year, which limits the variation in calf numbers, and places an upper limit on the effects of density-dependent calf survival. The models differ in terms of whether the population projections start when substantial catches first occurred or whether allowance is made for time-varying carrying capacity by starting the model in a more recent year. Most of the models are deterministic, but account needs to be taken of variation in cohort strength for analyses that include age-composition data or for species that are relatively short-lived. A limited number of analyses include process variability using a state-space-like modelling framework. For some stocks, abundance is so low that ideally both demographic and environmental variability should be included in models. The primary source of data for parameter estimation is a time-series of estimates of absolute abundance, although some approaches considered used a variety of data types, including relative abundance indices, mark-recapture data and minimum abundance estimates based on haplotype counts. In general, at least one estimate of absolute abundance is needed for parameter estimation; this is because there is a lack of catch-induced declines in abundance captured by indices of relative abundance that could be used to provide information on absolute abundance. Similarly, information on abundance from age- and length- composition data is limited. Most of the analyses quantify uncertainty using Bayesian methods to allow information on biological parameters, particularly the intrinsic rate of growth and the relative population at which maximum production occurs, to be included in the analyses, along with sensitivity testing. The future for the models on which assessments and management strategy evaluations are based will often involve multi-stock models that include age-, sex- and spatial-structure and are fitted as state-space formulations, although at present such models are often too computationally intensive to be feasible for implementation or there is insufficient information in the data to estimate the parameters representing all the processes, leading to simplifications, with the result that the performance of some of the methods of assessment used for cetacean stocks needs to be better understood, including through simulation testing.

KEYWORDS: ASSESSMENT; BAYESIAN; CETACEAN; MANAGEMENT STRATEGY

INTRODUCTION

Assessments of cetacean stocks[1] for use in management have, for several decades, been based on population dynamics models fitted to monitoring data. While conceptually similar to the approaches used to assess fish (Maunder and Punt, 2013) and invertebrate species (Punt *et al*., 2013), the assessment methods for cetacean stocks differ in some important ways. Specifically, catches (at least during the most recent three decades) have tended to be relatively low for most cetacean stocks – generally only bycatch, and in a few instances commercial, special permit or aboriginal subsistence catches. Therefore, information on absolute abundance provided by catch-induced declines in indices of relative abundance is not available. Consequently, most model-based assessments for cetacean stocks rely more on indices of absolute abundance than do assessments of fish and invertebrates. In addition, sample sizes for the age- and size-composition of removals are rarely high compared to those for commercially-important fish and invertebrate stocks.

Cetacean population assessments are used for a variety of purposes, often within specific contexts such as the conservation and management requirements of international bodies (such as the International Whaling Commission (IWC) or the North Atlantic Marine Mammal Commission (NAMMCO)) or national legislation (e.g. the US Marine Mammal Protection Act). In a conservation and management context, assessments can be used to provide:

(1) information on abundance in absolute terms and relative to the pre-exploitation size and to target and threshold levels;

(2) estimates of recent trends in abundance and/or mortality; and

(3) probabilities of rebuilding and extinction.

In addition, models are used to addresses questions of primarily scientific or ecological interest, such as which factors are driving population dynamics and distribution, and how might stocks or individuals respond to environmental change.

Particularly in the case of large whales in the context of the IWC, management advice is based on the application of 'management strategies'[2]. However, in most of the cases where catch limits (or 'strike' limits as they are called in the case of aboriginal subsistence whaling) are set for cetacean stocks, these are based on management strategies that use survey-based estimates of abundance, empirical rules that

---

[1] Stocks for the purposes of this review are generally taken to be management units. However, there is usually an attempt to use various sources of data to identify demographically independent units within a species or ocean basin e.g. see discussion in Donovan (1991).

[2] Combinations of data collection schemes, analysis methods and harvest control rules that have been selected using simulations that have evaluated their ability to achieve the management goals (Punt *et al*., 2016). Pioneered to a great extent by the IWC Scientific Committee, they are called 'management procedures' in the IWC and other cetacean literature e.g. the IWC's Revised Management Procedure (RMP) used to specify catch limits for baleen whales that are caught on their feeding grounds by commercial whalers and its Aboriginal Whaling Management Procedure (AWMP), which incorporates several stock-specific *Strike Limit Algorithms* to specify strike limits for subsistence hunts.

use survey estimates of abundance, or (in rare cases) simple model-based assessment methods combined with a harvest control rule. The selection of a management strategy should be based on simulation testing; a core element of simulation testing is the population dynamics model that represents the truth for the simulations (called the 'operating model'). The operating model is not an assessment model *per se*, but has many of the features of an assessment model and can be used to provide many of the types of outputs typically produced by an assessment. Thus, this review includes population models that have formed the basis for operating models, as well as those used to provide traditional outputs from stock assessments. For this reason, the term 'analysis' is used for the process of analysing monitoring data using methods that rely on some form of population dynamics model. However, and where appropriate, the term 'assessment' will be used to refer to a conventional stock assessment and 'MSE' to management strategy evaluation.

The aim of this paper is first to provide a summary of the stocks for which analyses have been undertaken and the analysis methods used most recently for those stocks. The analyses primarily concern large baleen whales in the context of the IWC and NAMMCO, which aim to conduct population model-based assessments, but some analyses exist for other stocks of cetaceans. The focus is on methods rather than results or whether those results were considered useful for management purposes (although in most cases, the assessments were approved by the relevant management bodies following a peer-review process). The results of this review are then used to identify (a) key issues that need to be addressed when deciding on the specifications for an analysis, (b) the choices made in current analyses and (c) the advantages and disadvantages associated with the different choices. Finally, the paper develops a set of 'best' practices for conducting cetacean stock assessments, with a focus on analyses for baleen whales (the subject of most analyses and for which most data are available).

STOCKS AND ANALYSES

The review focuses on recent (generally since 1995) analysis methods that involve population dynamics models that were applied to cetacean stocks. It does not cover the models used to analyse the monitoring data that provide the estimates of abundance for the assessments (e.g. Gerrodotte and Forcada, 2005; Cañadas *et al*., 2006), the models used to standardise catch-per-unit effort data (e.g. Cooke, 1993) or the models used to analyse mark-recapture data (e.g. Cooch and White, 2007)[3]. This review is restricted to analyses in which at least some of the parameters of the population dynamics model were estimated by fitting it to available data. Thus, model-based analyses in which all of the parameters are based on literature values/or 'guesstimated' by experts (e.g. Alvarez-Flores, 2006; Dueck and Richard, 2008; Reeves and Brownell, 2009; Slooten, 2015) are not covered in this review. Similarly, models that were based on back-calculation (e.g. Smith and Polacheck, 1979; Smith, 1983; Wade, 1993) where the only estimable parameter is carrying capacity and it is selected so that an estimate of current abundance is 'hit' are not the focus on this paper, nor are models that were developed primarily to estimate life history parameters (e.g. Fifas *et al*., 1998; Stolen and Barlow, 2003).

The stocks and hence the analysis methods summarised in this review were identified through: a literature search (Web of Science/Google Scholar); contacts with representatives of key management bodies; and contacts with individual analysts. Many of the reports describing analyses are found in the literature of management bodies (either the reports of the Scientific Committees of those bodies or papers presented to them) and are not necessarily searchable in databases such as Web of Science.

The results for 'large whales' (i.e. baleen whales and the sperm whale) are presented separately from those for other cetacean species, primarily because the peer-review process for analyses for baleen and sperm whales takes place through the IWC Scientific Committee, while that for the other species occurs as part of national (or in the case of some of the species harvested off West Greenland, the NAMMCO) review processes. The information is summarised by ocean basin or by stock, depending on the unit of analysis. In some cases, a stock has been assessed both as a single unit and as part of a regional analysis. In such cases, results are presented separately for the single unit and the regional analyses.

Tables 1 and 2 lists the stocks/species considered in this review, their major purpose (to form the basis of an assessment or to be the operating model for an MSE), the basic structure and some key (usually the most recent) references. However, in many cases the assessments were developed over several years. For example, Butterworth *et al*. (1999) outline an approach based on ADAPT-VPA for assessing Antarctic minke whales that was superseded by the integrated catch-at-age analysis method of Punt *et al.* (2014). Tables 3 and 4 outline the data types that were used in each analysis, while Tables 5 and 6 summarise how the analyses treated density-dependence, natural mortality and selectivity, three of the key processes that need to be included in any model-based analysis of a cetacean population. Finally, Tables 7 and 8 outline the types of outputs provided for each application and how uncertainty was quantified.

MODEL STRUCTURE ASSUMPTIONS

Population dynamics models

The assessments in Tables 1 and 2 are based on several types of population dynamics model. At the simplest level, are the analyses that aim only to estimate trends in abundance by fitting exponential models, perhaps using state-space formulations, to time-series of estimates of absolute abundance (e.g. those for eastern tropical Pacific spinner dolphins, and spotted dolphins). Such analyses provide no information about the status of stocks relative to reference points such as carrying capacity, but may provide information as to whether populations are increasing or not.

Most of the analyses in Tables 1 and 2 are based on age-structured models (often age- and sex-structured models) or production models. In general, the production models are based on the Pella-Tomlinson production function so that the point at which maximum surplus production occurs (MSYL, the Maximum Sustainable Yield Level) can be set to a value other than 0.5, with many assessments assuming that MSYL = 0.6. A small fraction of the population dynamics models also include stage structure. For example, Hoyle and

[3] Except where such data are integrated into an assessment model (e.g. Müller *et al*., 2011; Cooke *et al*., 2003, 2016).

Table 1

Assessments for baleen and sperm whales. '(a, b, etc.)' are used to distinguish among multiple analyses for the same species in this and all subsequent tables.

Species; stock	Purpose	Model types	Stock/spatial structure	Key references
BALAENOIDEA				
Bowhead whale				
Bering-Chukchi-Beaufort	Assessment (a, b); MSE (c, d)	Sex- and age-structured (a, b, c) Production model (d)	1 stock (a, b, c); 2 stocks (d)	Brandon and Wade (2006) (a); Punt (2006) (b); IWC (2003; 2008a) (c, d)
West Greenland	MSE	Sex- and age-structured	1 stock	IWC (2014a)
Southern right whale				
Southeast Atlantic	Assessment	Sex- and- and stage-structured	1 stock	Brandão *et al*. (2013)
Southwest Atlantic	Assessment	Sex- and- and stage-structured	1 stock	Cooke *et al*. (2003)
New Zealand	Assessment	Production model	1 stock	Jackson *et al*. (2016)
BALAENOPTERIDAE				
Common minke whale				
North Atlantic				
Eastern North Atlantic	Assessment	Sex- and age-structured	1 stock	Cooke (1993)
All	MSE	Sex- and age-structured	1, 2 or 3 stocks; 11 sub-areas	IWC (2017)
Western North Pacific	MSE	Sex- and age-structured	2, 3 or 5 stocks; 22 sub-areas	IWC (2014b)
Antarctic minke whales				
Indian and Pacific	Assessment	Sex- and age-structured	2 stocks (5 areas)	Punt *et al*. (2014)
Sei whale				
North Pacific	Assessment*			Tillman (1977)
Southern Hemisphere	Assessment*			Horwood (1980); Jones (1980)
Bryde's whale				
Western North Pacific	MSE	Sex- and age-structured	1 or 2 stocks (some with sub-stocks); two sub-areas	IWC (2008b)
Blue whale				
Eastern North Pacific	Assessment	Production model	1 stock	Monnahan *et al*. (2015)
Fin whale				
North Atlantic				
East Greenland-Iceland	Assessment	Sex- and age-structured	1 stock; sensitivity explored to alternative spatial definitions of the stock	Butterworth and Punt (1992)
All	MSE	Sex- and age-structured	3 or 4 stocks (some with sub-stocks); seven sub-areas	IWC (2017)
Southern Hemisphere	Assessment*			Breiwick (1977)
Humpback whale				
North Atlantic				
West Greenland	MSE	Sex- and age-structured	1 stock	IWC (2014a)
All	Assessment	Sex- and age-structured	2 stocks; 7 sub-areas	Punt *et al*. (2006)
Southern Hemisphere				
BSA (Brazil)	Assessment	Production model	1 stock	Zerbini *et al*. (2011)
BSB (West Africa)	Assessment	Production model	2–3 stocks off west Africa	Muller *et al*. (2011)
BSC (East Africa)	Assessment	Production model	2 stocks off east Africa	Johnston and Butterworth (2010)
BSD (W Australia) + BSE (E Australia) + BSF (Oceania)	Assessment	Production model	3 stock model	Ross Gillespie *et al*. (2014; 2015); Jackson *et al*. (2016)
BSG (West South America)	Assessment	Production model	1 stock	Johnston *et al*. (2011)

Cont.

Table 1 (*continued*).

Species; stock	Purpose	Model types	Stock/spatial structure	Key references
Gray whale				
North Pacific				
Western gray whale	Assessment	Individual-based stage-structured (26 stages) model	1 stock	Cooke *et al*. (2016)
Eastern gray whale	Assessment (a, b, c); MSE (d)	Production model (a); sex- and age-structured (b, d); sex-, age- and stage-structured (c)	1 stock	Wade (2002) (a); Punt and Wade (2012) (b); Brandon and Punt (2013) (c); IWC (2005) (d)
PCFG	MSE		2 stocks; no explicit spatial structure	IWC (2013)
All	Assessment	Sex- and age-structured	1 or 2 stocks (some with sub-stocks); 13 sub-areas	Punt (2016)
Sperm				
Western North Pacific	Assessment	Sex- and age-structured	1 stock; 12 fleets accounted for	Beddington and Cooke (1981)
		Sex-, age- and size-structured	1 stock; 12 fleets accounted for	Beddington and Cooke (1981); Cooke *et al.* (1983); Shirakihara, and Tanaka (1983); de la Mare and Cooke (1984)

*Out of date; not considered further.

Table 2

Assessments for other cetacean stocks. ‘(a)’, ‘(b)’, etc. are used to distinguish among multiple analyses for the same species in this and all subsequent tables.

Species; stock	Purpose	Model types	Stock/spatial structure	Key references
DELPHINOIDEA				
Hector’s dolphin				
Banks Peninsula	Assessment	Age- and area-structured	1 stock; four areas	Davies *et al*. (2008)
Spotted dolphin				
Eastern	Assessment	Exponential model (a); production model (a, b, d, e); age-sex-, and stage-structured (c)	1 stock	Wade *et al*. (2002) (a); Wade *et al*. (2007) (b); Hoyle and Maunder (2004) (c); Lennert-Cody *et al*. (2012) (d); Anon (2009) (e)
Southern/western	Assessment	Production model	1 stock	Anon (2009)
Spinner dolphin				
Eastern	Assessment	Exponential model (a); production model (a, b, c, d); age-structured model (a)	1 stock	Wade *et al*. (2002) (a); Wade *et al*. (2007) (b); Lennert-Cody *et al*. (2012) (c); Anon (2009) (d)
Whitebelly		Production model	1 stock	Anon (2009)
Common bottlenose dolphin				
Northern Adriatic	Assessment	Stochastic production model	1 stock	Simeoni (2014)
White whales				
Eastern Hudson Bay	Assessment	Exponential model	1 stock	Hammill *et al*. (2009)
Cook Inlet	Assessment	Stochastic sex- and age-structured	1 stock	Hobbs and Sheldon (2008); Hobbs *et al*. (2016)
West Greenland	Assessment	Sex- and age-structured	1 stock	Witting and Heide-Jørgensen (2015a)
Narwhal				
Northern Hudson Bay	Assessment	Production model	1 stock	Kingsley *et al*. (2012)
East Canada – West Greenland	Assessment	Sex- and age-structured	Eight areas (separately)	Witting (2015)
East Greenland	Assessment	Sex- and age-structured	Two areas (separately)	Witting and Heide-Jørgensen (2015a)
Harbour porpoise				
West Greenland	Assessment	Sex- and age-structured	1 stock	Witting (2013)
False killer whales				
Hawaii	Assessment	Production model	1 stock	Oleson *et al*. (2010)

Table 3
Summary of the data used in assessments of baleen whale stock assessments.

Species; stock	Catch	Absolute abundance	Relative abundance	Age/size structure	Other
Bowhead whale					
Bering-Chukchi-Beaufort	Yes	Yes	No	Age- and size-composition (b)	Proportion of calves and mature animals (a, b, c); assumptions about stock mixing rates (d)
West Greenland	Yes	Yes	No	No	
Southern right whale					
Southeast Atlantic	No	No	No	No	Yes
Southwest Atlantic	No	No	No	No	Yes
New Zealand	Yes	No	No	No	Yes
Common minke whale					
North Atlantic					
Eastern North Atlantic	Yes	Yes	CPUE	No	
All	Yes	Yes	No	No	Sex-ratio data (pre-fishery and current)
Western North Pacific	Yes	Yes	No	No	Bycatch estimates; mixing proportions; minimum and maximum abundance estimates
Antarctic minke whale					
Indian and Pacific	Yes	Yes	Yes	Age- and size-structure data from Japan	
Bryde's whale					
Western North Pacific	Yes	Yes	No	No	Mark-recapture
Blue whale					
Eastern North Pacific	Yes	Yes	No	No	Data on ship strikes
Fin whales					
North Atlantic					
East Greenland-Iceland	Yes	Yes	Yes (CPUE; account taken of correlations between CPUE indices for different years)	No	
All	Yes	Yes	Yes (CPUE account taken of correlations between CPUE indices for different years*)	Age-composition data for catches off Iceland	Mark-recapture
Humpback whale					
North Atlantic					
West Greenland	Yes	Yes	No	No	
All	Yes	Yes	Yes	No	
Southern Hemisphere					
BSA	Yes	Yes	Yes	No	Minimum abundance estimate based on haplotype counts
BSB	Yes	Yes	No	No	Mark-recapture data, minimum abundance estimate based on haplotype counts
BSC	Yes	Yes	Yes	No	Mark-recapture data
BSD + BSE + BSF	Yes	Yes	No	No	Mark-recapture data, minimum abundance estimate based on haplotype counts
BSG	Yes	Yes	Yes	No	Minimum abundance estimate based on haplotype counts
Gray whale					
North Pacific					
Western gray whale	No	No	No	Indirectly through individual hetero-geneity/ annual variation of pregnancy rates	Photo-ID; genetic sex determination
Eastern gray whale	Yes	Yes	No	No	Calf counts (b, c); Strandings and ice-cover (c)
PCFG	Yes	Yes	No		
All	Yes	Yes	No	No	Mixing proportions; bycatch numbers
Sperm					
Western North Pacific	Yes	No	No	Male length-frequency	Account is taken of male limitation on pregnancy rate

*Sensitivity test only

Table 4

Summary of the data used in assessments of other cetacean stocks.

Species; stock	Catch	Absolute abundance	Relative abundance	Age/size structure	Other
DELPHINOIDEA					
Hector's dolphin					
Banks Peninsula	Yes	Yes	Yes	Yes	Total survival (from mark-recapture); bycatch (assumed Poisson distributed); ages at first reproduction
Spotted dolphin					
Eastern	Yes	Yes	Yes (a) No (b,c,d,e)	No (a,b,d,e) Yes (c)	No
Southern/western	Yes	Yes	No	No	No
Spinner dolphin					
Eastern	Yes	Yes	Yes (a) No (b,c,d)	No	No
Whitebelly	Yes	Yes	No	No	No
Common bottlenose dolphin					
Northern Adriatic	Yes	Yes	No	No	No
Beluga whales					
Eastern Hudson Bay	Yes	Yes	No	No	No
Cook Inlet	Yes	Yes	No	Yes (by stage)	No
West Greenland	Yes	Yes	Yes	No	No
Narwhal					
Northern Hudson Bay	Yes	Yes	No	No	No
East Canada – West Greenland	Yes	Yes	No	No	No
East Greenland	Yes	Yes	No	Yes	No
Harbour porpoise	Yes	Yes	Yes	No	No
West Greenland	Yes	Yes	No	Yes	No
False killer whales					
Hawaii	No	Yes	Yes	No	No

Maunder (2004) represented the population of eastern tropical Pacific spotted dolphins using a model that kept track of age, sex and colour pattern. The more common use of stages in cetacean assessment models is to account for calving intervals that exceed a year. For example, the models developed by Brandon and Punt (2013) and Cooke *et al.* (2016) for gray whales and by Cooke *et al.* (2003) and Brandão *et al.* (2013) for right whales were stage-structured. Some of the assessments of sperm whales conducted by the Scientific Committee of the IWC were based on population dynamics models that tracked numbers of animals by sex and size-class.

The assessments of right whales in the southwest and southeast Atlantic (Cooke *et al.*, 2003; Brandão *et al.*, 2013) and of gray whales off Sakhalin Island (Cooke *et al.*, 2016) are examples of integrated mark-recapture-population dynamics models. The values for the parameters of the models on which these analyses were based were estimated by fitting the population model to the recapture histories for naturally marked animals. A key feature of these analyses is that data on newly-identified calves were used to provide information on calving rates and calving intervals. Unlike most of the models on which the analyses considered in this review are based (with the exception of the assessment of eastern North Pacific gray whales by Brandon and Punt, 2013), the models on which the assessments for southeast and southwest Atlantic right whales and gray whales off Sakhalin Island are based on dividing females into 'receptive', 'resting' and 'calving' classes to better mimic calving intervals. These analysis methods can be computationally intensive, especially if the aim is to quantify uncertainty using bootstrap and/or Bayesian methods, so their application has been limited to small populations (<1,000 animals in total) for which resighting probabilities are at least 10%.

Most of the analyses considered are for a single stock in a single area. However, there is an increasing trend towards accounting for spatial structure explicitly and including multiple stocks that mix and (in a limited number of cases) between which dispersal occurs. Many of these models were developed to form the basis for MSEs given the well-known sensitivity of the performance of management strategies for cetaceans to stock structure uncertainty (Punt and Donovan, 2007). Spatial and multi-stock models have been developed for bowhead, gray, common minke and humpback whales to account for the situation of feeding grounds where it is likely that animals from multiple stocks mix and for which there is no objective way to assign either catches or estimates of abundance to individual stocks. Another reason for including multiple stocks in analyses is when there are discrete feeding grounds, but the relationships among the animals on these grounds is unknown (e.g. Müller *et al.*, 2011, who identified ten alternative models/stock structure hypotheses for humpback whales off the west coast of Africa).

Many of the models on which cetacean assessments are based assume that stocks were at carrying capacity prior to exploitation and that carrying capacity has not changed over time. However, evidence for stocks such as the eastern North Pacific gray whales (Reilly, 1981, 1992; Cooke, 1986; Butterworth *et al.*, 2002) and humpback whales in the North Atlantic (Punt *et al.*, 2006) is that either carrying capacity has changed over time or some other assumptions of the model are badly violated (e.g. catches are greatly underestimated for example if struck and lost rates are markedly

Table 5

Technical aspects of the assessments of baleen and sperm whales.

Species; stock	Density-dependent component	Natural mortality	Selectivity (time-varying; estimated)	Other key features/notes
Bowhead whale				
Bering-Chukchi-Beaufort	None (a); births (Pella-Tomlinson form) (a, b, c); estimated MSYR (a, b); stochastic recruitment (c)[&]	Age-specific[$] (estimated) (a, b, c); episodic mortality events (c)[&]	Pre-specified (a, c, d); estimated (b)	Brandon and Wade (2006) examined several alternative models
West Greenland	Births (Pella-Tomlinson form); pre-specified MSYR (stochastic recruitment[&])	Pre-specified/episodic mortality events[&]	Pre-specified	
Southern right whale				
Southeast Atlantic	No	Age-specific[$] (estimated)	N/A	Time-variation variation in calving probabilities
Southwest Atlantic	No	Age-specific[$] (estimated)	N/A	Time-variation variation in calving probabilities
New Zealand	Pella-Tomlinson form (MSYR estimated with prior)	N/A	Implicitly assumed to be same as maturation	
Common minke whale				
North Atlantic				
Eastern North Atlantic	Births (Pella-Tomlinson form); estimated MSYR	Pre-specified (age-specific)	Pre-specified	
All	Births (Pella-Tomlinson form); pre-specified MSYR	Pre-specified (age-specific)	Pre-specified	
Western North Pacific	Births (Pella-Tomlinson form); pre-specified MSYR	Pre-specified (age-specific)	Pre-specified	Diffusion between sub-stocks (estimated); mixing rates (estimated); structured by season as well as space
Antarctic minke whale				
Indian & Pacific	Births (Ricker form); estimated MSYR, carrying capacity and recruitment deviates	Estimated (age-specific)	Logistic or dome-shaped by fleet; commercial selectivity time-varying	Random deviations in recruitment, mixing proportions; carrying capacity; growth
Bryde's whale				
Western North Pacific	Births (Pella-Tomlinson form); pre-specified MSYR	Pre-specified (age-specific)*	Pre-specified	Mixing rates (age-specific[&]) estimated
Blue whale				
Eastern North Pacific	Pella-Tomlinson form; estimated MSYR	N/A (production models)	Implicitly assumed to be same as maturation	Ship 'effort' predicted from model of shipping numbers
Fin whales				
North Atlantic				
East Greenland-Iceland	Births (Pella-Tomlinson form); estimated MSYR	Pre-specified*	Pre-specified*	
All	Births (Pella-Tomlinson form); pre-specified MSYR	Pre-specified	Estimated (logistic and time-invariant; logistic and time-dependent[&])	Diffusion between sub-stocks (estimated); mixing rates (estimated)
Humpback whale				
North Atlantic				
All	Births (Pella-Tomlinson/ Ricker form); estimated MSYR	Pre-specified	Pre-specified	Allowance is made for changing carrying capacity
West Greenland	Births (Pella-Tomlinson form); pre-specified MSYR (stochastic recruitment[&])	Pre-specified/episodic mortality events[&]	Pre-specified	
Southern Hemisphere				
BSA	Pella-Tomlinson form (MSYR estimated with prior)	N/A	Implicitly assumed to be same as maturation	
BSB	Pella-Tomlinson form (MSYR estimated with prior)	N/A	Implicitly assumed to be same as maturation	
BSC	Pella-Tomlinson form (MSYR estimated with prior)	N/A	Implicitly assumed to be same as maturation	Based on the 'sabbatical model'
BSD + BSE + BSF	Pella-Tomlinson form (MSYR estimated with prior)	N/A	Implicitly assumed to be same as maturation	
BSG	Pella-Tomlinson form (MSYR estimated with prior)	N/A	Implicitly assumed to be same as maturation	
Gray whale				
North Pacific				
Western gray whale	Individual heterogeneity / annual variation in pregnancy rate	Estimated		Accounts for the possibility of immigration; individual variation in sampling probability
Eastern gray whale	Births (Pella-Tomlinson form); estimated MSYR (a, b); estimated MSYR with stochastic recruitment (c); pre-specified MSYR (stochastic recruitment[&] (d)	Pre-specified/episodic mortality events[&]	Pre-specified	Effect of mass mortality event (estimated) (b, c); recruitment and natural mortality deviations linked to ice-cover (c)
PCFG	Births (Pella-Tomlinson form); pre-specified MSYR (stochastic recruitment[&])	Pre-specified/episodic mortality events[&]	Pre-specified	Diffusion (estimated); Effect of mass mortality event (estimated)
All	Births (Pella-Tomlinson form) pre-specified MSYR	Pre-specified	Pre-specified	Effect of mass mortality event (estimated); mixing rates (estimated)
Sperm				
Western North Pacific	Births (Pella-Tomlinson form); estimates MSYR and MSYL	Estimated; female linked to male	Tuned	

[$] Separately for calves and non-calves. *Sensitivity explored to alternative assumptions. [&] Sensitivity test only.

Table 6

Technical aspects of the assessments of other stocks.

Species; stock	Density-dependent component	Natural mortality	Selectivity (time-varying; estimated)	Other key features/notes
DELPHINOIDEA				
Hector's dolphin				
Banks Peninsula	Linear, with stochastic variation	Estimated (age-dependent; 5-parameter models)	Double-normal (estimated)	The model tracks abundance spatially, but the model itself is spatially-aggregated
Spotted dolphin				
Eastern	Pella-Tomlinson form (MSYR estimated) (a ,b, d, e); none (a); Pella-Tomlinson form with stochastic recruitment (MSYR estimated) (c)	Estimated (a, b, c, d, e); N/A (a)	Logistic (estimated) (c); implicitly assumed to be same as maturation (a, b, d, e)	Models in which r and K changed, with the year of change estimated were considered (a, b, e); Probability of moving between stages was assumed to logistic (c)
Southern/western	Pella-Tomlinson form (MSYR estimated)	N/A (production models)	Implicitly assumed to be same as maturation	Models in which r and K changed, with the year of change estimated were considered
Spinner dolphin				
Eastern	Pella-Tomlinson form (MSYR estimated) (a, b, c); None (a)	Estimated (a, b, c, d); N/A (a)	Implicitly assumed to be same as maturation (c)	Models in which r and K changed, with the year of change estimated were considered (a, b)
Whitebelly	Pella-Tomlinson form (MSYR estimated)	N/A (production models)	Implicitly assumed to be same as maturation	Models in which r and K changed, with the year of change estimated were considered
Common bottlenose dolphin				
Northern Adriatic	Schaefer model (MSYR estimated)	N/A (production model)	Implicitly assumed to be same as maturation	Carrying capacity allowed to change with prey abundance
White whales				
Eastern Hudson Bay	None	Included in intrinsic rate of growth	Pre-specified (uniform on 1+ animals)	State-space model; extent of struck and lost estimated; process variance pre-specified
Cook Inlet	Births and survival (Pella-Tomlinson form) (parameterised as growth rate); stochastic birth-death processes	Calculated from the growth rate	Pre-specified (uniform on mature animals)	Several stochastic processes; allowance is made for catastrophic mortality and Allee effects; K specified
West Greenland	Births (Pella-Tomlinson form) (MSYR estimated)	Estimated (age-specific)	Pre-specified (uniform on 1+ animals)	
Narwhal				
Northern Hudson Bay	None	Included in intrinsic rate of growth	Included in intrinsic rate of growth	State-space model; extent of struck and lost estimated; process variance pre-specified; includes killer whale predation
East Canada-West Greenland	Births (Pella-Tomlinson form) (MSYR estimated)	Estimated (age-specific)	Pre-specified (uniform on 1+ animals)	
East Greenland	Exponential growth (productivity estimated)	Estimated (age-specific)	Estimated	
Harbour porpoise				
West Greenland	Births (Pella-Tomlinson form) and exponential growth (productivity estimated)	Estimated (age-specific)	Estimated	Assessment not used for management
False killer whales	None	Implicit in the rate of growth	Not applicable	Scenarios explored two rates of change; Allee effects; catastrophic mortality event
Hawaii				

Table 7

Projection options, output statistics, and quantification of uncertainty for the assessments of baleen and sperm whales. Unless otherwise stated, the assessment examined uncertainty using sensitivity testing.

Species; stock	Projection ability	Reference points	Model output	Quantification of uncertainty
Bowhead whale				
Bering-Chukchi-Beaufort	No (a); Yes (b, c, d)	K, MSY, MSYR (a, b); K, MSYL (c, d)	RY, N^{tot} N^{tot}, need satisfaction, recovery rate, catch variation (b, c, d)	Bayesian
Southern right whale				
Southeast Atlantic	No	No	N^{mat}	Sensitivity analysis
Southwest Atlantic	No	No	N^{mat}	Sensitivity analysis
New Zealand	Yes	K	N, N/K	Bayesian
Common minke whale				
North Atlantic				
Eastern North Atlantic	No	None	N^{tot}, N^{exp}, N^{ma}	Asymptotic
All	MSE	K	N^{mat}/K^{mat}; predicted catches	Bootstrap
Western North Pacific	MSE	K	N^{mat}/K^{mat}; predicted catches	Bootstrap
Antarctic minke whale				
Indian and Pacific	Government of Japan (2016)	K, MSYR	Recruitment, N^{tot}, N^{tot}/K^{tot}	Asymptotic; likelihood profile*
Bryde's whale				
Western North Pacific	MSE	K	N^{mat}/K^{mat}; predicted catches	Bootstrap
Blue whale				
Eastern North Pacific	Yes	K	N, N/K	Bayesian
Fin whale				
North Atlantic				
East Greenland-Iceland	None	MSY, MSYR	RY, N^{tot}, N^{exp}, N^{tot}/K^{tot}, N^{exp}/K^{exp}	Bootstrap
All	MSE	K	N^{mat}/K^{mat}; predicted catches	Bootstrap
Humpback whale				
North Atlantic				
West Greenland	MSE	K	N^{tot}/K^{tot}; recovery rate, need satisfaction; catch variation	Bayesian
All	No	K, change in K		
Southern Hemisphere				
BSA	Yes	K	N, N/K	Bayesian
BSB	Yes	K	N, N/K	Bayesian
BSC	Yes	K	N, N/K	Bayesian
BSD + BSE + BSF	Yes	K	N, N/K	Bayesian
BSG	Yes	K	N, N/K	Bayesian
Gray whale				
North Pacific				
Western gray whale	Yes		N^{tot}, N^{mat}, Immigrants	Bayesian
Eastern gray whale	Yes (a, c, d)	MSY,RY (a); K (b, c, d)	N^{tot} (a,b); N^{tot}, need satisfaction (c); N^{tot}/K^{tot}; recovery rate, need satisfaction; catch variation (d)	Bayesian
PCFG	Yes	K	N^{tot}/K^{tot}; recovery rate, need satisfaction; catch variation	Bayesian
All	Yes	K	N^{tot}/K^{tot}; recovery rate, need satisfaction; catch variation	Bootstrap
Sperm				
Western North Pacific	No	K	N^{exp}/K^{exp}, N^{exp}	Sum of squares; likelihood profile

*Sensitivity test only.

Table 8

Projection options, output statistics, and quantification of uncertainty for other stocks.

Species; stock	Projection ability	Reference points	Model output	Quantification of uncertainty
DELPHINOIDEA				
Hector's dolphin				
Banks Peninsula	Yes	K	N^{ot}	Bayesian
Spotted dolphin				
Eastern	Yes (c); No (a, b, d, e)	K	N^{ot}	Bayesian (a, b, e); none (c)
Southern/western	No	K	N^{ot}	Bayesian
Spinner dolphin				
Eastern	No	K	N^{ot}	Bayesian
Whitebelly	No	K	N^{ot}	Bayesian
Common bottlenose dolphin				
Northern Adriatic	No	K	N^{ot}	Bayesian
White whales				
Eastern Hudson Bay	Yes (catches)	None	N^{tot}	Bayesian
Cook Inlet	Yes	Extinction	N^{tot}	Bayesian
West Greenland	Yes (catches)	RY, K	N^{tot}	Bayesian
Narwhal				
Northern Hudson Bay	Yes (catches)	None	N^{tot}	Bayesian
East Canada-West Greenland	Yes (catches)	K	N^{tot}	Bayesian
East Greenland	Yes (catches)	K	N^{tot}	Bayesian
Harbour porpoise				
West Greenland	Yes (catches)	RY, K	N^{tot}	Bayesian
False killer whales				
Hawaii	Yes	Near extinction	N^{tot}	Bayesian

higher than believed). In this respect, the Bering-Chukchi-Beaufort Seas stock of bowhead provides an illuminating example. Earlier assessments of this stock (e.g. Givens *et al.*, 2005) were able to fit the available data under the assumption of time-invariant carrying capacity. However, the most recent data indicate that the rate of increase has not slowed down as would be expected for a population that is approaching its carrying capacity. Consequently, the most recent models for this stock of bowhead whales (e.g. Punt, 2015a) do not assume that carrying capacity has been constant for 150 years and instead, following Wade (2002), start the population projections in 1940, with the age-structure at that time assumed to be stable. Punt and Butterworth (2002) started population projections for the Bering-Chukchi-Beaufort Seas stock of bowhead whales from various years and assumed that the age-structure at that time corresponded to a population increasing an estimated rate.

In general, there is little need to include multiple fleets in model-based analyses for cetaceans, unlike the case for fish and invertebrates where differences in catch age- or size-compositions among areas or groups of vessels are often addressed by assuming that fishery selectivity differs spatially or seasonally. This is because it is difficult for whalers to select for animals of particular ages and even sizes. In addition, catch data are usually available by sex. However, spatial variation in age structure may interact with the spatial distribution of the whaling operations to produce apparent spatial and temporal differences in selectivity. There are some analyses with multiple fleets e.g. for the eastern North Pacific stock of gray whales and common minke whales off West Greenland, owing to differences in selectivity patterns between commercial and aboriginal whalers. Multiple fleets were also considered in the assessments for sperm whales in the North Pacific as a proxy for spatial structuring of the population, and in the assessments of Antarctic minke whales. The latter assessment allows for time-varying commercial selectivity given among-year changes in where the various fisheries operated.

Density-dependence

Density-dependence may operate on a variety of population processes including maturation, growth, calving rate, juvenile survival, adult survival and even movement rates. However, it is seldom the case that sufficient data are available to estimate the parameters governing even one of these processes.

The models that assume that population size has been increasing exponentially have no explicit representation of density-dependence. Brandon and Wade (2006) compared several models for the Bering-Chukchi-Beaufort Seas stock of bowhead whales and found that the highest posterior probability was assigned to the model that did not start the population projections when catches were first recorded and ignored density-dependence[4]. The analyses based on mark recapture data only (i.e. those for gray whales off Sakhalin Island and right whales in the southwest and southeast Atlantic) also do not account for density-dependence. These populations are all assessed to be increasing exponentially so any estimates of density-dependence parameters (and carrying capacity) would be very uncertain anyway.

With one exception, assessments that allow for density-dependence assume that it operates on births, generally assuming the Pella-Tomlinson density-dependence form i.e. the expected number of calves during year y, C_y, is given by:

$$C_y = N_y^m f_0(1 + A(1 - (N_y^d/K^d)^2)) \tag{1}$$

---

[4] This conclusion was strengthened once additional abundance data were collected (Punt, 2015a).

where N_y^m is the number of females capable of calving during year y, f_0 is the pregnancy rate at carrying capacity, A is the resilience parameter, z is the degree of compensation, N_y^d is the magnitude of the density-dependence component of the population during year y, and K^d is the magnitude of the density-dependence component of the population at carrying capacity. The parameter z is related to the value of MSYL, while the value of A is related to both the maximum pregnancy rate and the Maximum Sustainable Yield Rate[5]. Punt (1999) provides the relationships among A, z, MSYL and MSYR for the case of an age- and sex-structured population dynamics model. Equation 1 can lead to negative numbers of calves when the population is larger than K, which is clearly unrealistic so the constraint is usually imposed that the number of calves cannot be less than zero. Such a constraint can lead to convergence problems when minimisation is based on software that requires a differentiable objective function (such as AD Model Builder, Fournier *et al*., 2012). Consequently, the assessment of Antarctic minke whales by Punt *et al*. (2014) assumed a Ricker-like formation of equation 1, which implies that the number of calves tends to zero for $N_y^d/K^d \gg 1$.

It is possible to assume that density-dependence acts on births (equivalent in most cases to density-dependence on fecundity or calf mortality) and/or non-calf survival (Punt, 2015b). However, only one of the analyses (that for Cook Inlet white whales[6], Hobbs and Sheldon, 2008; Hobbs *et al*., 2016) included density-dependent natural mortality.

Other population dynamic assumptions

The base versions of the analyses are generally similar, but there are often many differences in the alternative models examined to explore uncertainty. The focus in this paper is on the assumptions for the base versions of the models. The key differences among the models are discussed below.

Is the population dynamics model deterministic or is some aspect of the dynamics stochastic?

The most general model in this respect is that developed for Antarctic minke whales that allows for deviations in (a) recruitment about the density-dependence function (i.e. about expected calf numbers), (b) the proportion of the population in each area in which two stocks are found, (c) selectivity spatially and over time and (d) in carrying capacity. Several other assessments (generally of shorter-lived species) consider stochastic recruitment, including the model developed Hoyle and Maunder (2004) for eastern tropical Pacific spotted dolphins, that for false killer whales, that for Cook Inlet (Alaska) white whales and that for Hector's dolphins off Banks Peninsula, New Zealand. Several of the analyses consider the possibility of episodic events in the future, but only the analyses for the eastern North Pacific gray whales estimate an episodic event (or catastrophe) in the past (for which direct evidence exists). Some stocks are 'very small', necessitating modelling of both demographic and environmental variation (e.g. Breiwick and Punt, 2002).

Is natural mortality (M) age-, sex- or stage-structured?

In general, the values for the parameters related to natural mortality or survival for cetaceans are pre-specified (Tables 5 and 6) and in some cases, natural mortality depends on age (e.g. for fin and common minke whales in the North Atlantic and North Pacific). Some of the analyses estimate natural mortality (and in the case of Antarctic minke whales how natural mortality depends on age). Hoyle and Maunder (2004) assumed there was an age-at-senescence, an assumption that was not made in other analyses (the limited evidence available suggests that senescence does not occur in baleen whales). Survival is, however, poorly estimated unless age data are available for which selectivity can either be estimated precisely of for which selectivity can reasonably be assumed to be uniform. Obtaining data to estimate juvenile and calf natural mortality is more challenging than for estimating adult natural mortality (if it is assumed that mortality is constant for all adult ages).

What is the first year of the modelled period?

Conventionally, analyses for cetacean stocks started in the first year for which (non-trivial) catches were recorded and it was assumed that the stock was at carrying capacity at that time. However, increasingly analyses are being conducted in which the model projections start after the stock has been subject to high previous catches. This is either because the earlier catches are considered to be very uncertain (or simply unknown) or because the assumption that the stock was at carrying capacity when catches were first recorded is incompatible with recent trends in estimates of abundance. In general, however, the estimates of carrying capacity from analyses in which the projections start fairly recently are imprecise. The exception is for stocks such as the eastern North Pacific stock of gray whales for which the rate of increase in abundance has declined, suggesting that the population is now approaching its (new) carrying capacity.

Has carrying capacity or productivity changed over time?

Most of the assessments assume that K and MSYR have remained constant over time. The assessments that start the population projections in a year more recently than when the first catches were recorded (e.g. Brandon and Wade, 2006), implicitly assume that carrying capacity may have changed over time (and, for the eastern North Pacific gray whales, models that assume time-invariant K are unable to mimic the trend in abundance inferred from the survey data) and some of the analyses for dolphins in the eastern tropical Pacific considered models in which K changed at some point in the past (with the year in which the change occurred treated as an estimable parameter). Thus, these analyses implicitly postulate that a regime shift in K occurred (without specifying the cause). The Antarctic minke whale assessment estimates changes over time in K as a random walk, thereby avoiding having to specify (or estimate) when it changed. Estimation of MSYR is challenging (e.g. IWC, 2014c) even when it is assumed to be time-invariant and thus consideration of time-varying productivity is unusual[7]. However, the analyses of dolphin populations in the eastern tropical Pacific considered model variants that estimated two

---

[5] MSYR, the ratio of MSY to the equilibrium number of recruited animals when the population is producing MSY.

[6] 'White whales' are also commonly known as 'beluga' or 'beluga whales'; the common name is white whale as agreed by the IWC Scientific Committee is used throughout this paper.

[7] Scenarios in which productivity is assumed to change over time are, however, commonly included in MSEs.

levels for MSYR (modelled as the intrinsic rate of growth), i.e. implicitly assuming that a regime change in productivity occurred.

How is selectivity modelled?

The choice of the fishery selectivity pattern is probably inconsequential when the catch is small relative to the population size and there are no data on the age- or size-composition of the catch. Consequently, many analyses based on age-structured models make simple assumptions regarding fishery selectivity, such as that selectivity is uniform above age 1 or selectivity is pre-specified based on historical assumptions (e.g. for common North Atlantic minke whales). However, the availability of age-composition data has allowed selectivity to be estimated for some stocks (Antarctic minke whales, North Atlantic fin whales, the Bering-Chukchi-Beaufort Seas stock of bowhead whales, sperm whales in the western North Pacific, spotted dolphins in the eastern tropical Pacific, and narwhals and harbour porpoise off West Greenland). The assessment of Antarctic minke whales in the southern Pacific and Indian Ocean appears to be the only assessment that explored alternative functional forms for selectivity (dome-shaped vs asymptotic). This exploration supported the use of sex-specific dome-shaped selectivity that changed over time and differed spatially. Dome-shaped and spatial differences in selectivity are probably a consequence of the spatial distribution of the population (larger animals tend to be closer to or in the ice and hence less available to the fleet), while selectivity would differ over time as a function of where in the large areas on which the model is based the fishery operated in. Correct specification of selectivity is particularly important when catch age- or length-composition data are used for parameter estimation because these data can have a large influence on estimates of absolute abundance unless they are highly down weighted. Misspecification of selectivity can lead to biased estimates of exploitation rate and hence abundance.

How are the values of biological parameters linked to environmental factors?

In principle, environmental drivers of the population dynamics can be represented implicitly by estimating parameters such as the annual deviations in calf numbers about those expected given the deterministic relationship between abundance and pregnancy rate. Only one assessment (Brandon and Punt, 2013) attempted to explicitly link an environmental variable (ice-cover) to the deviations in calf numbers. Cooke (2007) outlined a modelling framework that does not explicitly model environmental drivers of productivity, but that allows a distinction to be made between the maximum growth rate a species can achieve in ideal habitat and the average growth rate that a population at low level will achieve in a given habitat. This model leads to the conclusion that the rate of population growth will be closer to deterministic at low stock size than close to carrying capacity. Cooke (2016) outlined how reference points such as MSYL can be defined when carrying capacity is varying over time.

The models that consider spatial structure rarely represent spatial structure explicitly, i.e. no attempt is made to define the probability that whales in one area move to another area. Rather, the models that consider spatial structure estimate (or pre-specify) the proportion of each stock in each area. The estimates of the mixing proportions are based primarily on data on the proportion of each stock in each area (e.g. using genetic information). In general, the models that include multiple stocks assume that there is no permanent transfer of animals between stocks (known as ‘diffusion’); exceptions include the models developed to test management strategies for common minke whales in the western North Pacific, fin whales in the North Atlantic and gray whales off the west coast of North America.

All but one of the analyses considered are based on models with an annual time-step. The exception is the model on which the MSE for the western North Pacific common minke whales is based, which operated on a monthly time-step to capture the impact of catching animals during their migration.

DATA USED FOR ASSESSMENT PURPOSES

The key data inputs to a stock assessment/MSE are a time-series of catches (ideally by fleet and sex), along with an index of relative or absolute abundance. The primary source on trends in abundance are estimates of abundance from surveys (Tables 3 and 4). Some earlier assessments (e.g. Cooke, 1993; Butterworth and Punt, 1992) were based on analyses of commercial catch and effort data. However, catch-rate-based indices of abundance are now considered to insufficiently reliable for use in cetacean assessments without a level of detail of knowledge of the operations that rarely, if ever, exists (IWC, 1989).

Catches were included in most of the analyses (Tables 3 and 4). However, catches, particularly those for the earliest years of exploitation (in some cases as early as the 16th century), often need to be adjusted by struck and lost rates (e.g. Smith and Reeves, 2003). Most analyses for baleen and sperm whales only considered removals due to modern (*ca*. post-1865) commercial and aboriginal harvesting, although more recently other direct removals such as bycatches in fishing gear and deaths due to ship strikes are being considered. For example, the model used for the rangewide assessment of Pacific gray whales (Punt, 2016) also included bycatch data, while that on which the assessment of eastern North Pacific blue whales was based included the impact of shipstrikes. In contrast to large whales, most of the anthropogenic removals of dolphins and porpoises are due to bycatch; estimating robust bycatch estimates is difficult and such estimates are usually much more uncertain than catches by commercial whaling (e.g. Lewison *et al.,* 2004).

All but one of the analyses considered made use of estimates of absolute abundance for parameter estimation purposes. A noteworthy exception is the models developed for sperm whales in the western North Pacific, which were fitted to the catch length-frequency data for males (although these data were subsequently found to be unreliable). Those models were developed in the early 1980s, prior to the start of most of the major abundance survey programmes. Consequently, should new assessments of western North Pacific sperm whales be undertaken, they would probably use survey estimates of abundance (perhaps as relative indices of abundance given difficulties estimating the number of animals missed for long-diving species such as sperm whales). In general, analyses that fit to data on trends in absolute abundance involve analysing data from sighting surveys to provide estimates of abundance that are then treated as data in a second analysis that estimates parameters such as productivity and carrying capacity. This is appropriate when the estimates of abundance are independent. However, this is not the case when sample sizes are small such that some parameters are assumed to be same among years. Moore and Barlow (2013) analyse survey data for beaked whales off the west coast of North America in which trend estimation is conducted simultaneously with abundance estimation and model changes in abundance using a deterministic exponential model – in principle changes in abundance could have been represented using a model in which annual changes in abundance were stochastic, i.e. using a full state-space model.

Several of the analyses also made use of data on relative abundance. These are usually estimates of abundance from surveys, but when it has not proven possible to estimate the constant of proportionality for the surveys, often because $g(0)$, the probability of detecting a school on the trackline, is not equal to 1 and cannot be estimated, or surveys only cover only a proportion of the area in which the stock being assessed is found. In the latter case, the estimates of relative abundance may be biased due to temporal variation of the proportion of the stock inside the survey area.

There was generally only a single estimate of absolute abundance for the earliest assessments that used such data for parameter estimation (e.g. Butterworth and Punt, 1992) and assessments selected the value for K such that the model 'hit' that abundance estimate (e.g. Smith and Polacheck, 1979; Smith, 1983; de la Mare, 1989). However, as additional surveys were conducted, it was possible to include the abundance data in the likelihood function maximised to estimate the values for the parameters. Increasing numbers of surveys led to the observation (e.g. Wade, 2002) that the sampling standard deviations for the survey estimates were too small given the demographics of cetaceans, i.e. the estimates varied more among years than was possible for a long-lived animal. This has led to now common practice of estimating an 'additional variance' parameter in analyses in which there are multiple estimates of absolute or relative abundance. This additional variation may represent sampling error, temporal variation in the constant of proportionality between survey estimates of abundance and actual abundance, unmodelled stochastic population dynamics or model misspecification.

Some methods for estimating abundance share parameters among years (e.g. Zeh and Punt, 2005; Laake *et al*., 2010), while other methods analyse sightings data pooled over several years (e.g. Bøthun and Øien, 2011). This leads to the error in the estimates of abundance being correlated, which must be accounted for in the likelihood function assumed for the estimates of abundance (e.g. Givens *et al*., 1995). The analyses for the eastern North Pacific gray whales and the Bering-Chukchi-Beaufort Seas stock of bowhead whales include a variance-covariance matrix for the estimates of absolute abundance.

Mark-recapture data are available for several stocks. These data have been used to estimate (a) mixing rates for North Atlantic fin whales and western North Pacific Bryde's whales (b) abundance for southwest and southeast Atlantic right whales, gray whales off Sakhalin Island and several Southern Hemisphere humpback whales and (c) survival for Hector's dolphins off Bank's Peninsula. Mark-recapture data are commonly used to estimate abundance, but several of the analyses for Southern Hemisphere humpback whales have instead integrated the mark-recapture data directly into the analysis (Table 3). Reasons for this include being able to account for losses in numbers due to natural mortality directly, as well as to let the data on trend from the mark-recapture data enter the analyses. In principle, mark-recapture data may imply a non-significant trend in abundance, but a statistically significant trend may be detected when all of the information for the stock is taken into account. Caution needs to be taken to ensure that the data are appropriately weighted when multiple sources of data are included in an analysis.

Several of the assessments of Southern Hemisphere humpback whale stocks included a constraint on the lower bound for the total number of animals in the population based on counts of mtDNA haplotypes[8]. To be included in an assessment in the form of a lower bound for the minimum total number of animals (N_{min}), the observed number of haplotypes needs to be corrected for sampling probability, for the number of males and the number of immature animals and for the number of haplotypes that might have been lost subsequent to the population being at its lowest level. In general, the impact of imposing an N_{min} is greatest when it is large because N_{min} places an implicit constraint on the maximum rate of increase (and hence MSYR).

Age- and size-composition data are only available for a small number of cetaceans and these are the species/stocks for which selectivity and deviations in calf numbers from expectation have been estimated. The data tend to be downweighted given a lack of independence in the sampling process, particular for commercial catches (e.g. Punt *et al*., 2014); such downweighting is common in assessments of fish and invertebrate stocks (e.g. McAllister and Ianelli, 1997; Francis, 2011; Punt, 2017). Care needs to be taken when including age- and length-composition data in analyses because while these data can provide information on absolute abundance, the information is sensitive to model misspecification, particularly misspecification of the selectivity function. Hobbs *et al.* (2016) fit their model to data on the proportion of the catch that consists of immature animals, mature females and mature males. Other data sources included in population analyses for cetaceans include the proportion of calves and mature animals from aerial surveys (Bering-Chukchi-Beaufort Seas bowhead whales), the sex-ratio of catches (North Atlantic minke whales), mixing proportions based on genetics data (eastern North Pacific gray whales, western North Pacific minke whales), and calf counts (eastern North Pacific gray whales).

MODEL FITTING AND QUANTIFICATION OF UNCERTAINTY

The models on which the analyses are based were, with a few (historical) exceptions, fitted using maximum likelihood or Bayesian methods.

---

[8] The observed number of haplotypes in a population provides an absolute minimum of the number of females when the population was at its lowest level (Jackson *et al*., 2006).

Measures of statistical uncertainty

Most of the analyses have attempted to quantify parameter uncertainty using Bayesian, bootstrap, or asymptotic methods (e.g. Wade, 1999; Tables 7 and 8), although other methods such as Monte Carlo methods and likelihood profiling have also been applied. The bootstrap approach has been used most extensively to quantify the uncertainty associated with values for the parameters of the operating models on which MSEs have been based. These operating models are usually based on pre-specifying the parameter that determines productivity (normally expressed as MSYR), which is usually poorly determined even in data rich situations (e.g. Punt *et al*., 2014; de la Mare, 2016). The bootstraps tend to be parametric, where data are generated from their sampling distributions, and the model fitted to each such bootstrap data set.

Most of the analyses in Tables 1 and 2 quantified uncertainty using Bayesian methods (Tables 7 and 8). There are a variety of reasons for this. One is simply historical precedence – some of the first uses of Bayesian methods in assessments of harvested marine populations were for cetaceans (e.g. Givens *et al*., 1995). In addition, production of posterior distributions is computationally feasible for many cetacean stocks given the relatively limited amount of data available. More importantly, given the limited amount of information contained in the data for most stocks (e.g. for MSYL), Bayesian methods provide a way to include prior information in analyses. Whilst priors can be assumed to be uniform (e.g. Wade *et al*. 2002, 2007), it is preferable to base a Bayesian analyses on ‘data-based’ priors that are informative and represent a synthesis of parameter estimates among species and stocks. Most of the analyses in Tables 1 and 2 based on Bayesian methods imposed priors on biological parameters such as age-at-maturity, maximum pregnancy rate and the survival rates for calves and non-calves (with the constraint imposed that the calf survival rate cannot exceed that of non-calves). Placing a prior on the maximum pregnancy rate is equivalent to imposing a prior on MSYR or the maximum growth rate. However, there is often little information to update the priors (e.g. the eastern North Pacific blue whales), and in some cases, priors are updated to biologically unrealistic or implausible values. Zerbini *et al*. (2010) used information about biological parameters, in conjunction with an age-structured model, to develop a probability distribution for the maximum rate of increase for humpback whales. Furthermore, IWC (2014b) used a Bayesian approach to construct a probability distribution for the rate of increase for whale stocks that were severely depleted when data collection started, and this distribution was used to select a minimum plausible bound for MSYR expressed in terms of the 1+ component for the population for use in MSEs for baleen whales. It is difficult to impose upper bounds on biological parameters such as survival rate, age-at-maturity and maximum pregnancy rate because these parameters tend to be highly correlated (Brandon *et al*., 2007).

The difficulties of specifying priors are well known. In the context of assessments of cetaceans, the key discussions have related to (a) whether it is reasonable to impose independent priors on each of age-at-maturity, survival rate and maximum pregnancy rate given observed correlations between the values for these parameters when estimates can be made, (b) which parameters to impose priors on, specifically because priors for parameters for which information is lacking are often assumed to be uniform (e.g. should a prior be imposed on MSYL or z, both of which relate to the shape of the production function), and (c) should a prior be imposed on K or abundance in a recent year[9]. In general, while data can update the prior for K (or current abundance) and perhaps productivity, parameters such as the age-at-maturity and MSYL are seldom updated much.

An important difference between assessments for fish and invertebrate populations and those for cetaceans is that for cetaceans, catches tend to be low compared to productivity, particularly during recent years when most of the monitoring data are available i.e. there is no information on absolute abundance from catch-induced changes in relative abundance. However, parameters related to the density-dependence function can be estimated when stocks were depleted prior to the collection of indices of relative and absolute abundance and the monitoring data cover a period during which the population was increasing at close to the maximum possible rate (IWC, 2015; Tables 3 and 4).

Sensitivity analyses

All but one of the analyses considered examined sensitivity to assumptions by changing some of the assumptions of the base model (or a set of base models). The exploration of sensitivity tends to be most extensive for MSEs since one of the primary aims is to identify a management strategy that is robust to uncertainty; in designing an MSE, the aim should be for the set of operating models to be reduced with additional research (Punt *et al*., 2016). The set of operating models must be reasonable so that selection of the management strategy is not dictated by unrealistic assumptions and thus it rare that even MSEs will explore *all* ‘plausible’ hypotheses and assumptions[10]. Nevertheless, the number of sensitivity tests can be substantial for some MSEs. Table 9 provides an example for a single stock situation and the number can be considerably greater where there is uncertainty regarding stock structure, which can involve changing the number of stocks in the region being managed and where they are located (e.g. fin and common minke whales in the North Atlantic and common minke whales in the western North Pacific).

Most of the sensitivity tests for assessments involve changing the values for pre-specified parameters, changing the priors imposed on the parameters as part of Bayesian analyses, and (much less often) considering different structural models and different functional forms for natural mortality and selectivity.

Simulation evaluation

It is now considered essential in resource management to evaluate the performance of assessment methods before they are used to provide management advice. The IWC Scientific

---

[9] Most Bayesian cetacean assessments now place a prior on current abundance to avoid the prior for carrying capacity being updated prior to inclusion of data simply because some combinations of productivity and carrying capacity are inconsistent with the population being currently extant given the model and historical catches.

[10] Assigning ‘plausibility’ remains one of the greatest challenges in developing and implementing management procedures within the IWC (Punt and Donovan, 2007).

Table 9

The *Robustness Trials* (sensitivity tests) for the MSE for the Bering-Chukchi-Beaufort Seas bowhead whales (from IWC, 2003).

Trial	Factor	Basic trials	Factor level
BR01	A: Density-dependence	1, 1^{S}, 9, 9^{S}	Density-dependence on mature (BE trials use 1+)
BR02	B1: Stochastic dynamics	8^{S}, 13^{S}	Stochastic dynamics (with serially-correlated environmental variation)
BR04	E: Survey frequency	9, 13, 14, 16, 20	(a) 15 yrs
		16, 20	(b) 5 yrs
BR05	F: Strategic surveys	9	(a) Yes + CV = (0.25, 0.25)
		9	(b) Yes + CV = (0.34, 0.25)
BR06	G: Survey bias time dependence	1	(a) Historic bias (1978–2002): 1.5 constant; Future bias: decreasing (1.5 → 1)
		1	(b) Historic bias (1978–2002): 0.67 constant; Future bias: increasing (0.67 → 1)
		9, 9^{S}	(c) Future bias: sinusoidal from base value in yr 0 to maximum of 150% in yr 40
		12, 12^{S}	(d) Future bias: decreasing (1.5 → 1) from year 0 to 100
		14	(e) Future bias: increasing from 1 → 1.5 in year 25 and constant thereafter (former BE15)
BR07	H: Future survey *CV*	1	(a) CV = (0.1, 0.1)
		1^{S}	(b) CV = (0.34, 0.25)
		9	(c) (0.1, 0.1) + sinusoidal survey bias
BR08	I: Historic catch bias	14, 16^{S}	(a) 0.5 bias from 1848–1914
		14, 16, 16^{S}	(b) 1.5 bias from 1848–1914
BR09	K: Time dependence in *K*	1, 9, 10	(a) *K* halves linearly over 100 years
		1, 9, 10	(b) *K* doubles linearly over 100 years
		1	(c) *K* sinusoidal from base value in year 0 to maximum of 150% in year 40
		1, 21	(d) Tent *K: K* doubles linearly from years-50 to 0 and halves from years 0 to 50
		1, 9	(e) *K* halves linearly over 100 years + strategic surveys
BR10	L: Time dependence in *MSYR*	10	(a) Resilience (*A*) halves linearly over 100 years,
		9	(b) Resilience (*A*) doubles linearly over 100 years
		1, 8	(c) Resilience steps 2½% → 1% → 2½% every 33 yrs over 100 years
		1, 8	(d) Resilience steps 2½% → 1% → 2½% every 33 yrs over 100 years in sync with *M* (compute MSYR first) – if it is practical halve *M* for each age class
		1	(e) *K* and *A* halve linearly over 100 years
		1, 21	(f) *K* and *A* vary as tent (see BR09)
BR11	M: Time dependence in *M*	1, 9, 10	(a) Natural mortality *M* halves linearly over 100 years (and calculate resulting MSYR)
		1, 9, 10	(b) *M* doubles linearly over 100 years
BR12	N: Episodic events	1, 1^{S}, 9, 9^{S}	2 events occur, between years 1–50, in which 20% of animals die
BR13	O: Integrated	1, 11, 14	(a) $MSYR_{1+} \sim U[0.01, 0.04]$; fixed $MSYL_{1+} = 0.6$
		11, 14	(b) $MSYR_{1+} \sim U[0.01, 0.04]$; $MSYL_{1+} \sim U[0.4, 0.8]$
		1, 1^{S}, 11, 11^{S}, 14	(c) $MSYR_{1+} \sim U[0.01, 0.04]$; $MSYL_{1+} \sim U[0.4, 0.8]$; historical catch bias ~ U[0.5, 1.5]; Serial correlation ~ U[0.47, 0.95]
		1^{S}, 11^{S}	(d) $MSYR_{1+} \sim U[0.01, 0.04]$; $MSYL_{1+} \sim U[0.4, 0.8]$; historical catch bias ~ U[0.5, 1.5]; Serial correlation $\rho \sim U[0.47, 0.95]$, time delay in density-dependence ~ U[0, 30]
BR14	P: 1st year of population projection	1, 9	1940 (reference or base case level is 1848 or 1748 for stochastic trials)
BR15	$MSYL_{1+} = 0.9$	1, 9, 10	
BR16	B2: Different stochastic parameter	1^{S}	(a) Negative correlation in recruitment $\rho = -0.75$
		1^{S}	(b) High correlation in recruitment $\rho = 0.9$
		1^{S}, 9^{S}, 10^{S}	(c) High correlation in recruitment $\rho = 0.9$; + Episodic events
		1^{S}	(d) Change σ_ε^2 to give 3* variation in population size at equilibrium
		9^{S}	(e) $\rho = 0.9$ + change σ_ε^2 to give 3*equilibrium variation + episodic events

[s] Stochastic.

Committee pioneered the testing of stock assessment methods using simulation (e.g. Kirkwood, 1981; de la Mare, 1986). For example, the estimation performance of the length-structured models used for assessment of sperm whale stocks in the western North Pacific was explored in several simulation studies (e.g. Cooke and de la Mare, 1983; Shirakihara and Tanaka 1984; Shirakihara *et al*., 1985; de la Mare, 1988).

In contrast to the situation for fisheries assessments (see the summary in Table 6 of Dichmont *et al.,* 2016), only a relatively small proportion of the methods on which the analyses in Table 1 and 2 are based have been subject to simulation evaluation. This is due in part to several of these methods being computationally extensive. However, there are some examples of recent assessment methods (including Bayesian methods) having been evaluated using (often limited) simulation including: (a) the Bering-Chukchi-Beaufort Seas stock of bowhead whales (Punt and Butterworth, 1997); (b) Antarctic minke whales in the Indian and Pacific Oceans (Punt and Polacheck, 2008; de la Mare, 2016); and humpback whales off the east and west coasts of Australia (Leaper *et al*., 2011).

PROJECTIONS AND MANAGEMENT OUTPUTS

Most, but not all, of the methods on which the analyses considered are based have the capability to project into the future (Tables 7 and 8). The models developed as the basis for operating models for MSEs and those that have formed the basis for Population Viability Analysis (PVA) are the most general in this respect. The assessments tend to be used to evaluate the implications of future series of catches, or simply to project the population ahead in the absence of exploitation to estimate the time for the population to reach some proportion of *K* or other target level. Perhaps the most extensive evaluation of the future state of a cetacean population was that of Hobbs *et al*. (2016) for white whales in Cook Inlet, Alaska. In addition to removals due to hunts, they considered the impact of predation by killer whales (in the past and in the future), catastrophic events in the future, as well as mass mortality events. They did not estimate posterior distributions for these processes, but rather examined sensitivity to alternative plausible values for the parameters governing them. The assessment of Antarctic minke whales reported time-trends in calf numbers, as well as growth rates and carrying capacity (this can only be done for the few assessments that estimate changes over time in recruitment, growth and carrying capacity).

In contrast to the assessments, the MSEs evaluate full-feedback management strategies. They thus include a component that generates the types of future data (usually absolute abundance data[11]) that will form the basis for new assessments. This contrasts with fisheries MSEs where it is common to generate several types of data including catch rate indices of relative abundance, catch age-and size-composition data, survey indices of abundance, along with the associated survey age- and size-composition data (Punt *et al*., 2016). The relative lack of data generated by cetacean MSEs reflects the data available for most species (Tables 3 and 4), and the fact that management strategies, even those based on population models (such as the IWC's RMP and AWMP) use relatively few data types. The MSEs generally assume that all of the removals are managed using the management strategy under evaluation, but there are some exceptions to this, including the IWC's strategy for the Greenlandic hunt for bowhead whales where account is taken of catches by Canada (which is not a member of the IWC). The evaluation of recent implementations of the IWC's RMP for commercial catches of fin and common minke whales in the central and western North Atlantic was based on MSEs that pre-specified the catches in aboriginal hunts from the same stocks (IWC, 2017).

The common outputs from analyses (and their projections) are time-trajectories of numbers of animals in absolute terms or relative to reference points such as K or MSYL. The population numbers are usually summarised as the total population size, although some assessments also report numbers of all females (e.g. Cooke *et al*., 2016) or mature females. Some of the earlier assessments for the eastern North Pacific stock of gray whales (e.g. Wade, 2002) and for the Bering-Chukchi-Beaufort Seas stock of bowhead whales (e.g. Givens *et al*., 1995; Brandon and Wade, 2006) reported estimates of current replacement yield (the catch so that the population size in the next year equals that at the start of the present year), as this quantity formed the basis for management advice before IWC *Strike Limit Algorithms* were developed for these stocks in 2005 and 2003 respectively.

The MSEs are capable of producing a large number of outputs. The most common outputs include the final depletion (the ratio of the mature population size at the end of the projection period to carrying capacity or the mature population size at the end of the projection period in the absence of exploitation had there been no catches – when carrying capacity is changing over time), the lowest depletion (or the ratio of the mature population size to that which would have arisen had there been no catches) over the projection period, and the recovery rate for depleted populations. The MSEs that have evaluated management strategies for commercial whaling have reported average catches as well as catch variation, and those that have evaluated management strategies for aboriginal subsistence whaling have reported what fraction of the need of aboriginal communities can be satisfied.

DISCUSSION

'Best' practices for modelling cetacean stocks

Table 10 lists a set of 'best practice' guidelines for conducting analyses for cetacean stocks. The words 'best practice' here are used in the way that has become common parlance but it should be recognised that what is important is that analyses *are adequate for the purpose they are intended* – 'best' in the sense of the 'best available at present' may not necessarily be adequate depending on what they are to be used for and 'adequate' analyses may not always need to be the best available That being said, the best practices suggested here represent (a) a synthesis of recent modelling decisions for cetacean stocks; (b) best practices in the field of fisheries assessment and wildlife modelling; and (c) highlight those factors that are likely to be consequential for the provision of management advice and should at least be considered in analyses. They are based primarily on the experience of the author – in principle, they could be tested using simulation studies, but this is beyond the scope of the present paper.

Choice of modelling structure

The type of model on which the analyses are based is determined in part by the sizes of the populations. The analyses that rely on mark-recapture data (e.g. those for gray whales off Sakhalin Island, and those for right whales in the Atlantic) are tailored to populations that are in the low 100s of animals. Nevertheless, some of the analyses based on age- and sex-structured population dynamics models, and population dynamics models that are sex- and age-aggregated have been applied to populations that are relatively small (e.g. low 100s Cook Inlet white whales and Banks Peninsula Hector's dolphins) as well as to populations consisting of thousands to hundreds of thousands of individuals (e.g. Antarctic minke whales, dolphin stocks off the west coast of North America).

The state of the art in terms of population projections for marine renewable resources is to allow for parameter uncertainty, and stochastic dynamics (demographic uncertainty as well as environmental stochasticity) in the future. Analyses of stocks in the low 100s of animals should ideally account for both demographic and environmental stochasticity. In contrast, analyses for large populations can safely ignore the effects of demographic uncertainty, but should still consider the impact of environmental stochasticity, particularly for birth rates and survival. The estimates of parameters related to environmental stochasticity may be very imprecise unless data on, for example, catch age-composition are available, which is uncommon for many cetacean stocks. Unlike fish and invertebrates, the number of calves-per-female is constrained for a cetacean. Consequently, there are limits to the amount

---

[11] Other data could include the proportion of the population that are calves, juveniles or adults (e.g. for the Bering-Chukchi-Beaufort Seas stock of bowhead whales; IWC, 2003).

Table 10

Tentative best practice guidelines for cetacean stock assessments.

Issue	Guidelines
Model structure assumptions	
Spatial and stock structure	Required if evidence suggests population structuring within the area being assessed or perhaps if there is limited information to assess possible stock structuring (the absence of information is not information on absence).
Age- and sex-structure	Should be the default (sex-structure can be ignored if demographic parameters do not differ between the sexes and the sex ratio of the historical removals is close to 1–1).
Stage-structure	Often unnecessary, but can be used to impose assumptions regarding calving intervals or where the data suggest it might be important.
First year of the model	Ideally, the first year for which catches are available so that population can be assumed to have been at carrying capacity at the start of the first year with removals. However, a later year may be appropriate if the historical removals are very uncertain, or carrying capacity is likely to have changed over time.
Demographic stochasticity	Not needed for populations of 1,000 or more animals.
Environmental stochasticity	Worth including in base-case models when there is evidence for catastrophic events or simply for stocks for which there is likely to be among-year variation in pregnancy rate and or calf survival. Should be considered routinely if data on age- or size-composition are available.
Key biological and fishery processes	
Density-dependent processes	Models should consider density-dependence in birth rate and adult natural mortality.
Natural mortality	Should be age- or stage-specific (minimally calf, non-calf; but alternative forms such as the Siler form should be considered).
Selectivity	Usually only required to be estimated if removals are a substantial proportion of the population or if age- or size-composition data are included in the likelihood function. In principle, selectivity should depend on fleet, and consideration should be given to domed-shaped and time-varying selectivity.
Time-varying parameters	These pertain to selectivity, growth, distribution, and calf mortality, and should be treated as random effects (with the extent of variation estimated).
Model fitting	
Additional variation	The presence of additional variance should be tested for and accounted for. Similarly, the extent of overdispersion should be estimated for age- and size-composition data to avoid overfitting these data.
Prior distributions	Consider, to the extent possible, the use of data-based priors, and place priors on current abundance rather than carrying capacity.
Fit to raw data rather than summarised data	Where possible, models should be fit to the data in their rawest form (e.g., recapture histories instead of estimates of abundance from program MARK) to avoid the methods for analysing the raw data and those underlying the population making different sets of assumptions.
Use a state-space formulation	Inclusion of time-varying parameters requires the specification of parameters that constrain the extent to which such parameters can vary over time. Sensitivity can be explored to the values for these parameters if they have to be pre-specified rather than being estimated (e.g. Punt *et al*., 2014).
Uncertainty quantification	
Primary basis for quantify uncertainty	Bayesian methods permit prior information to be included in analyses and produce the information needed for the basis for projections (the probability associated with alternative parameter vectors and even models).
Sensitivity tests	These should be as broad as possible, ideally divided into 'more plausible' and 'less plausible' sets.
Simulation evaluation	Test the performance of the estimation method using simulations prior to their application in a management context.

by which the number of calves can differ from the expected value given by equations such as equation 1. Punt *et al*. (2014) recognised this and formulated the function defining recruitment variation to impose an upper bound on the numbers of calves-per-female in any year and the model of Cooke (2007) shows that large reductions in population size can occur on an annual basis, but this is not the case for increases owing to population demography. In general, stochasticity in calf numbers has limited impact on population trajectories when calf survival is larger than 0.9. However, such stochasticity must be modelled if the model is to be fitted to age-, size- or stage-composition data or if calf survival is to be linked to an environmental variable such as ice cover.

The choice between using a production model and an age- and sex-structured population dynamics model is semi-arbitrary although analyses for stocks with age-, size- or stage-composition data would logically be based on models that have this type of structure. Nevertheless, the choice between basing an analysis on an age-structured population dynamics model or a production model is often a pragmatic (computational) one, especially when the aim is to quantify uncertainty using Bayesian methods, there are multiple stocks of the species of interest in the region, or there is a substantial amount of informative data.

Experience has shown that there is often little justification for the inclusion of sex-structure in analyses. However, it is prudent to explicitly model sex-structure for species for which the catch sex-ratio can be markedly different from 1:1 (such as Antarctic and common minke whales and gray whales), because the relative reduction of the two sexes could differ markedly. Obviously, the number of calves will be directly related to the number of mature females, but social behaviour related to reproduction might result in the number of males also markedly impacting reproduction rates, particularly in odontocetes.

Most early analyses assumed that the region under consideration contained only a single stock. However, the available data (including mark-recapture, telemetry and genetics data) often suggest that multiple stocks of a given species may be found in a region, and these stocks may mix where catches and surveys occur. In such cases, it is necessary to develop multi-stock population dynamics models. The present models that allow for multiple areas and movement do not model movement explicitly but rather treat the proportions of each stock in each modelled area as estimable parameters (or pre-specify these parameters); this seems reasonable.

Finally, most models ignore within-year dynamics. This is reasonable for cetaceans, which are long-lived and for which removals are generally a small proportion of total abundance. The operating model developed for the western North Pacific common minke whales was the only one that allowed for seasonal dynamics; this was needed because catches occur during migration, and consequently the stock-, sex-, and age-composition of the catches in some areas changes during the year.

In general, estimation performance (measured by the precision with which parameters such as *K* is estimated) is improved if the stock is assumed to be at *K* at the start of the first year for which substantial catches are available. However, the benefits of improved estimation ability may be lost if the historical catches are subject to considerable uncertainty or if there are regime shifts in carrying capacity. In such cases, it may not be possible to provide reasonable estimates of population size relative to reference points such as carrying capacity and MSYL.

Parameterisation of processes

Most of analyses for cetacean stocks are based on models that represent the age- and sex-structure of the population (the analyses for Southern Hemisphere humpback whales being a notable exception). Age- and sex-structured models require specifications for how density-dependence is represented, as well as how survival, maturity and fishery selectivity are modelled as a function of age or sex.

Most past analyses have assumed that density-dependence impacts calf survival/fecundity/age-at-maturity (the effects of which tend to be difficult to distinguish) and whilst it is reasonable for this to remain the default, assuming density-dependence in adult survival can lead to different population dynamics so this source of density-dependence is worth at least considering in analyses.

Natural mortality is probably age- (or at least stage) specific. This can be modelled by assuming that calf survival differs from that for non-calf animals (assuming that calf survival is the square of adult survival is a simple way to force this to be true). However, if there are age-composition data, it may be possible to model age-specific natural mortality using a functional form such as the Siler model (Siler, 1979). Moore and Read (2008) used age-composition data to fit the Siler model for harbour porpoises. Punt *et al*. (2014) considered the Siler model as well as that natural mortality changes as an auto-regressive process with age, but eventually selected a piecewise linear model for natural mortality-at-age with breakpoints based on the results of initial model runs, for simplicity.

The way selectivity is modelled will be largely inconsequential owing to the longevity of most cetaceans. However, selectivity should be estimated rather than being pre-specified if (a) historical removals were very large and (b) if age- (or size-) composition data are included in the likelihood function. This is because composition data can provide information on absolute abundance, but such estimates are sensitive to misspecification of selectivity. In general, it is reasonable to assume that selectivity is an asymptotic function of age or size. However, this assumption should be tested if there are fleets for which the assumption that selectivity is asymptotic is likely to be invalid, and there are data for those fleets.

In principle, parameters for natural mortality, growth, selectivity, *K* and distribution could be linked (perhaps with error to environmental variables [Brandon and Punt, 2013]). However, selecting the correct variables can be challenging. Thus, in general, it is better to treat parameters that may vary over time as random effects, possibly (as in Brandon and Punt, 2013) linked to an environmental variable.

Main sources of uncertainty/quantification of uncertainty

Assessments for cetacean species and stocks are subject to a wide variety of sources of uncertainty. The major source of uncertainty is likely to be stock-specific. Punt *et al*. (2016) identify the categories of uncertainty that should be considered for inclusion in the operating models on which MSEs are based. The uncertainties that usually have the greatest impact on estimates of current abundance, and current abundance relative to reference points are: (a) model structure uncertainty, in particular in the context of analyses of cetaceans, uncertainty about stock structure (number of stocks, where they are found, how they move, and whether there is permanent movement among them); (b) uncertainty about the constant of proportionality between estimates of abundance and abundance itself; and (c) uncertainty about historical catches (particularly if these are large relative to sustainable yields). The performance of management strategies usually depends on the uncertainties that impact estimation of current abundance, but also on uncertainties related to (a) the quality and frequency of future data, and (b) regime shifts in productivity, natural mortality and carrying capacity.

Care needs to be taken to ensure that the way the uncertainties are characterised is plausible. This is particularly the case for uncertainties that relate to possible future events (e.g. future changes in carrying capacity and productivity, an increased frequency of episodic events), as current data may not shed much light on the likelihood of such events. Butterworth *et al*. (1996) outline a scheme for evaluating the relative plausibility of alternative hypotheses that could form the basis for sensitivity analyses in MSEs. In general, as is the case for recent MSEs undertaken by the IWC, it is advisable to divide sensitivity tests into a reference set (called ‘Evaluation’ trials in IWC parlance) that consists of the more plausible sets of assumptions and a less plausible set (called ‘Robustness’ trials in IWC parlance) that includes scenarios that are of interest to examine the behaviour of the management strategies in more ‘extreme’ circumstances.

Several ways have been used to quantify uncertainty (Tables 7 and 8), but the trend for cetacean assessments is towards the use of Bayesian approaches, notwithstanding the challenges associated with specifying defensible prior distributions. This is because (a) Bayesian methods permit the inclusion of prior information, in particular about the intrinsic rate of growth (or equivalently the MSYR) and (b) because the outputs of a Bayesian analysis are the inputs for decision analysis (i.e. the probability of alternative parameter vectors and even alternative models).

Data

In general, it is better to consider using as many sources of data as possible in assessments, although as mentioned earlier catch per unit of effort data are unlikely to be representative of changes in population size. However, it must be recognised that model misspecification, including incorrect assumptions about sampling error, can degrade results when multiple data sources are used for parameter estimation. In addition, inclusion of multiple data sources, can lead to identification of data conflicts, and hence the need to weight different data sources. In general, it is advisable to follow the recommendation of Francis (2011) that assessments should always try to mimic the trends in the index of abundance best, if they are representative of the stock, perhaps at the expense of fits to age-composition data. Age- (or size-) composition data are required if selectivity (or natural mortality) is to be estimated (although given the demographics of whales, the value for adult survival can often be informed by the rate of increase).

Simulation testing

Although many of the earlier methods of assessment for cetacean stocks were subject to evaluation using simulation, the use of simulation to evaluate estimation methods is now less common that was the case 20–30 years ago. While this perhaps reflects the complexity of some of the estimation methods, it is not good practice and is counter to the improved trend in fisheries assessment where most of the key methods have been subject to some form of simulation evaluation (e.g. Dichmont *et al*., 2016).

Some future directions

It is beyond the scope of this paper to provide a full analysis of improvements in assessments methods and MSEs that are or might be forthcoming. A short, annotated list of some key areas is presented below.

(1) Multi-species modelling to provide management advice

The analyses outlined in the paper ignore biological interactions among species – multi-species modelling is certainly an avenue to be explored but for a number of reasons is not yet at the stage of being able to provide robust management advice (e.g. Mori and Butterworth, 2006; Schweder *et al*., 1998).

(2) Use of individual-based models

In principle, it is possible to apply individual-based models to cetacean assessments and MSEs, especially for ‘small’ populations. In effect, this is close to the approach of the existing mark-recapture-based assessments for southwest and southeast Atlantic right whales and Sakhalin Island gray whales. Punt and Breiwick (2002) outlined an assessment and MSE framework that is based on an individual-based population dynamics model. This framework was developed to evaluate management strategies for small stocks, but has not been used to date.

(3) Incorporating non-lethal and cumulative effects data

Few analyses explicitly address the issue of the cumulative impacts of non-lethal impacts (such as the impact of the Gulf of Mexico oil spill on bottlenosed dolphins and the reduction in salmon numbers on killer whale survival in Puget Sound). In large part, this is because of the lack of available data (and hence understanding) of non-lethal factors individually and cumulatively to assess their impact on cetacean dynamics. Efforts are underway with respect to chemical pollutants (e.g. Hall *et al.,* 2016), noise and other forms of disturbance (e.g. King *et al.,* 2015), and food availability (de la Mare, 2017). The latter model is individual-based and has been used to better understand population-level yield curves.

(4) Incorporating ‘raw’ data in assessments

Most assessments fit the population model to estimates of abundance when these are determined from surveys. Recently, there has been a move to include data sources in assessments in their raw form (e.g. the integration of mark-recapture histories directly into the population models for right whales in the southwest and southeast Atlantic and for gray whales off Sakhalin Island). Nadeem *et al*. (2016) outline an approach in which raw sightings data for fin whales off the US west coast are fitted within a state-space population dynamics model. The state-space model used is based on an age- and sex-aggregated model, with production based on a Gompertz model and no allowance for historical removals. In principle, this approach could be extended to account for age, sex and catches, but this might come at a substantial computational cost.

FINAL THOUGHTS

Model-based assessments of cetaceans, especially baleen whales, remain the gold standard for providing management advice. Assessments for cetaceans usually have (and rely on) at least one estimate of absolute abundance. This is stark contrast to fisheries assessments where absolute abundance must be inferred from changes in relative abundance and age-composition. That being said, estimation of trends in abundance (and hence the values for parameters such as MSYR) in fishery assessments rely on information such as trends in relative abundance or age-composition, the latter of which is rare for cetacean stocks.

The ability to estimate stock status relative to reference points such as carrying capacity or MSYL for cetaceans is challenging in those cases where the catch history is long and uncertain and/or carrying capacity may have changed since the start of substantial catches. This issue is also a concern for fisheries (e.g. those in Europe and the east coast of North America) where exploitation started many centuries before the establishment of monitoring programmes. However, this problem can be partially overcome for these fisheries given the availability of often substantial amounts of catch and survey age-composition data during periods when exploitation rates and biomass changed substantially.

This review has shown that there are generally fewer data available for parameter estimation purposes for cetaceans than for fish and invertebrates (although better independent estimates of absolute abundance). Fisheries science has much to learn from analyses conducted for cetaceans, in particular the way MSE has been applied, the use of data independent of commercial catches, and the attempts to better understand/evaluate the implications of alternative stock structure hypotheses.

ACKNOWLEDGEMENTS

This work was conducted under contract to the Inter-American Tropical Tuna Commission. Mark Maunder, Cleridy Lennert (IATTC), Malcolm Haddon (CSIRO and Oceans and Atmosphere), Justin Cooke (CEMS), and two anonymous reviewers are thanked for comments on early versions of this paper. Greg Donovan (IWC Secretariat) improved the manuscript with his usual 'few minor editorial suggestions'.

REFERENCES

Alvarez-Flores, C.M. 2006. A review of the analysis to evaluate the recovery of bowhead whales, *Balaena mysticetus*, in the eastern Canadian Arctic. DFO Canadian Science Advisory. Research Document 2006/054. iv+36pp.

Anon. 2009. Updated estimate of NMIN and stock mortality limits. Document SAB-07-05 presented to the 7th Meeting of the Scientific Advisory Board of the International Dolphin Conservation Program, 30 October 2009, La Jolla, CA. 12pp. [Available from: *https://www.iattc.org/PDFFiles2/SAB-07-05-Nmin-and-Stock-Mortality-Limits.pdf*]

Beddington, J.R. and Cooke, J.G. 1981. Development of an assessment technique for male sperm whales based on the use of length data from the catches, with special reference to the North-west Pacific stock. *Rep. int. Whal. Commn* 31: 747–60.

Bøthun, G. and Øien, N. 2011. Report of the Scientific Committee. Annex D. Report of the Sub-Committee on the Revised Management Procedure. Appendix 7. Estimates used for catch limit calculations in northeast Atlantic minke whales. *J. Cetacean Res. Manage.* (Suppl.) 12: 107–10.

Brandão, A., Butterworth, D.S., Ross-Gillespie, A. and Best, P.B. 2013. Application of a photo-identification based assessment model to southern right whales in South African waters, now including data up to 2012. Paper SC/65a/BRG17 presented to the IWC Scientific Committee, June 2013, Jeju Island, Republic of Korea (unpublished). 15pp. [Paper available from the Office of this Journal].

Brandon, J.R. and Punt, A.E. 2013. Testing the Gray Whale Strike Limit Algorithm (SLA): allowing environmental variability to influence population dynamics. *J. Cetacean Res. Manage.* 13(1): 81–8.

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. Manage*. 8(3): 225–39.

Brandon, J.R., Breiwick, J.M., Punt, A.E. and Wade, P.R..2007. Constructing a coherent joint prior while respecting biological bounds: application to marine mammal stock assessments. *ICES J. Mar. Sci.* 64: 1,085–1,100.

Breiwick, J.M. 1977. Analysis of the Antactic fin whale stock in Area I. *Rep. int. Whal. Commn* 27: 124–27.

Butterworth, D.S. and Punt, A.E. 1992. Assessments of the East Greenland-Iceland fin whale stock. *Rep. int. Whal. Commn* 42: 671–96.

Butterworth, D.S. Korrubel, J.L. and Punt, A.E. 2002. What is needed to make a simple density-dependent response population model consistent with data for eastern North Pacific gray whales? *J. Cetacean Res. Manage.* 4(1): 63–76.

Butterworth, D.S., Punt, A.E. and Smith, A.D.M. 1996. On plausible hypotheses and their weighting, with implications for selection between variants of the Revised Management Procedure. *Rep. int. Whal. Commn* 46: 637–40.

Butterworth, D.S., Punt, A.E., Geromont, H.F., Kato, H. and Fujise, Y. 1999. Inferences on the dynamics of Southern Hemisphere minke whales from ADAPT analyses of catch-at-age information. *J. Cetacean Res. Manage.* 1(1): 11–32.

Cañadas, A., Fortuna, C. and Hammond, P.S. 2006. Modelling techniques to investigate the impact of changes in habitat on cetacean distribution and abundance. Paper SC/58/E21 presented to the IWC Scientific Committee, May 2006, St. Kitts and Nevis, West Indies (unpublished). 11pp. [Paper available from the Office of this Journal].

Cooch, E. and White, G. 2007. *Program MARK: A Gentle Introduction.* 900pp. [Available at *www.phidot.org/software/mark/docs/book/*].

Cooke, J.G. 1986. On the net recruitment rate of gray whales with reference to inter-specific comparisons. *Rep. int. Whal. Commn* 36: 363–6.

Cooke, J.G. 1993. Further analyses of minke whale CPUE in the Northeast Atlantic, with estimates of yield rates. Paper SC/45/NA7 presented to the IWC Scientific Committee, April 1993 (unpublished). 18pp. [Paper available from the Office of this Journal]

Cooke, J.G. 2007. The influence of environmental variability on baleen whale sustainable yield curves. Paper SC/59/RMP10 presented to the IWC Scientific Committee, May 2007, Anchorage, USA (unpublished). 15pp. [Paper available from the Office of this Journal]

Cooke, J.G. 2016. A note on the selection of stock recruitment relationships for western Atlantic Bluefin tuna, with reference to temporal variation in the relationship. *Collect. Vol. Sci. Pap. ICCAT.* 72: 1,449–56.

Cooke, J.G. and de la Mare, W.K. 1983. Description of and simulation studies on the length-specific sperm whale assessment technique. *Rep. int. Whal. Commn* 33: 741–6.

Cooke, J., Rowntree, V. and Payne, R. 2003. Analysis of inter-annual variation in reproductive success of South Atlantic right whales (*Eubalaena australis*) from photo-identification of calving females observed off Península Valdés, Argentina, during 1971–2000. Paper SC/55/O23 presented to the IWC Scientific Committee, May 2003, Berlin (unpublished). 16pp. [Paper available from the Office of this Journal].

Cooke, J.G., de la Mare, W.K. and Beddington, J.R. 1983. An extension of the sperm whale model for the simulation of the male population by length and age. *Rep. int. Whal. Commn* 33: 731–4.

Cooke, J.G., Weller, D.W., Bradford, A.L., Sychenk, O., Burdin, A.M., Lang, A.R. and Brownell, R.L., Jr. 2016. Updated population assessment of the Sakhalin gray whale aggregation based on a photoidentification study at Piltun, Sakhalin, 1995–2015. Paper SC/66b/BRG225 presented to the IWC Scientific Committee, Bled, Slovenia, June 2016 (unpublished). 13pp. [Paper available from the Office of this Journal]

Davies, N., Bian, R., Starr, P., Lallemand, P., Gilbert, D. and Mckenzie, J. 2008. Risk analysis of Hector's dolphin and Maui's dolphin subpopulations to commercial set net fishing using a temporal-spatial age-structured model. Final Research Report to the Ministry of Fisheries, Wellington, NZ. 113pp.

De la Mare, W.K. 1986. Fitting population models to time series of abundance data. *Rep. int. Whal. Commn* 36: 399–418.

De la Mare, W.K. 1988. Extended sensitivity trials on the length-specific method for estimating the size of exploited sperm whale populations. *Rep. int. Whal. Commn* 38: 243–8.

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.

De la Mare, W.K. 2016. Simulation studies on the properties of estimates from statistical catch at age models. Paper SC/66b/IA8 presented to the IWC Scientific Committee, Bled, Slovenia, June 2016. 34pp. Paper available from the Office of this Journal]

De la Mare, W.K. 2017. Further developments of the individual based energetics model and progress on the development of an emulation model. Paper SC/67a/RMP2 presented to the IWC Scientific Committee, Bled, Slovenia, May 2017 (unpublished). [Paper available from the Office of this Journal]

De la Mare, W.K. and Cooke, J.G. 1984. A new model for the female sperm whale incorporating aspects of the reproductive cycle. *Rep. int. Whal. Commn* 34: 251–4.

Dichmont, C.M., Deng, R., Punt, A.E., Brodziak, J., Chang, Y-J, Cope, J.M., Ianelli, J.N., Legault, C.M., Methot, R.D., Porch, C.E., Prager, M.H. and Shertzer, K. 2016. A review of stock assessment packages in the United States. *Fisheries Research* 183: 447–60.

Donovan, G. P. 1991. A review of IWC stock boundaries. *Rep. int. Whal. Commn* (Special Issue 13): 39–68.

Dueck, L. and Richard, P. 2008. Preliminary evaluation of selective hunting and recovery of Eastern Canadian Arctic-West Greenland bowhead whales. Canadian Science Advisory Secretariat Research Document 2008/088. 17pp.

Fifas, S., Goujon, M. and Antoine, L. 1998. Application of Leslie's model on a population of common dolphns (*Delphinus delphis*): Senstivity study. *Aquat. Liv. Res.* 11: 359–69.

Fournier, D.A., Skaug, H.G., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M.N., Nielsen, A., and Sibert, J. 2012) AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. *Optimization Methods and Software* 27: 233–49.

Francis, R.I.C.C. 2011. Data weighting in statistical fisheries stock assessment models. *Can. J. Fish. Aquat. Sci.* 68: 1,124–38.

Gerrodette, T. and Forcada, J. 2005. Non-recovery of two spotted and spinner dolphin populations in the eastern tropical Pacific Ocean. *Marine Ecology Progress Series* 291: 1–12.

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.

Government of Japan. 2016. Results of the analytical work on NEWREP-A recommendations on sample size and relevance of age information for the RMP. Paper SC/66b/SP10 presented to the IWC Scientific Committee, Bled, Slovenia, June 2016 (unpublished). 23pp. [Paper available from the Office of this Journal]

Hall, A.J., Schwacke, L.H. and Rowles, T.K. 2016. Assessing the effects of petroleum-associated compound exposure on cetacean populations using the Pollution 2020 individual based modelling approach. Paper SC/66b/E04 presented to the IWC Scientific Committee, June 2016, Bled, Slovenia (unpublished). 10pp. [Paper available from the Office of this Journal]

Hammill, M.O., Kingsley, M.C.S, Lesage, V. and Gosselin, J-F. 2009. Abundance of Eastern Hudson Bay beluga. DFO Canadian Science Advisory. Secretariat Research Document. 2009/009. iv+18 p.

Horwood, J.W. 1980. Population biology and stock assessment of Southern Hemisphere sei whales. *Rep. int. Whal. Commn* 30: 519–30.

Hobbs, R.C. and Shelden, K.E.W. 2008. Supplemental status review and extinction assessment of Cook Inlet belugas (*Delphinapterus leucas*). AFSC Processed Rep. 2008-08, Alaska Fish. Sci. Cent., NOAA, NMFS, 7600 Sand Point Way NE, Seattle WA 98115. 76pp.

Hobbs, R.C, Wade, P.R. and Sheldon, K.W. 2016. Viability of a small, geographically-isolated population of beluga whales, *Delphinapterus leucas*: Effects of hunting, predation, and mortality events in Cook Inlet, Alaska. *Mar. Fish. Rev.* 77: 59–88.

Hoyle, S.D. and Maunder, M.N. 2004. A Bayesian integrated population dynamics model to analyse data for protected species. *Animal Biodiversity and Conservation* 27: 247–66.

International Whaling Commission. 1989. Report of the Comprehensive Assessment Workshop on Catch per Unit Effort (CPUE), Reykjavik, 16–20 March. *Rep. int. Whal. Commn* (Special Issue 11): 15–28.

International Whaling Commission. 2003. Report of Standing Working Group on the Aboriginal Subsistence Whaling Management Procedure. *J. Cetacean Res. Manage.* (Suppl.) 5: 154–225.

International Whaling Commission. 2005. Report of Standing Working Group on the Aboriginal Subsistence Whaling Management Procedure. *J. Cetacean Res. Manage.* (Suppl.) 7: 115–87.

International Whaling Commission. 2008a. Report of Standing Working Group on the Aboriginal Subsistence Whaling Management Procedure. *J. Cetacean Res. Manage.* (Suppl.) 8: 121–49.

International Whaling Commission. 2008b. Report of Sub-Committee on the Revised Management Procedure. *J. Cetacean Res. Manage.* (Suppl.) 10: 90–119.

International Whaling Commission. 2012. The Revised Management Procedure (RMP) for Baleen Whales. *J. Cetacean Res. Manage.* (Suppl.) 13: 485–94.

International Whaling Commission. 2013. Report of the Standing Working Group on the Aboriginal Subsistence Whaling Management Procedure. *J. Cetacean Res. Manage*. (Suppl.) 13: 130–53.

International Whaling Commission. 2014a. Report of the Standing Working Group on the Aboriginal Subsistence Whaling Management Procedure. *J. Cetacean Res. Manage.* (Suppl.) 15: 189–213.

International Whaling Commission. 2014b. Report of Sub-Committee on the Revised Management Procedure. *J. Cetacean Res. Manage.* (Suppl.) 15: 87–188.

International Whaling Commission. 2014c. Report of the Fourth Intersessional Workshop on the Review of Maximum Sustainable Yield Rates (MSYR) in Baleen Whales, 26–28 March 2013, La Jolla, California, USA. *J. Cetacean Res. Manage*. (Suppl.) 15: 507–518.

International Whaling Commission. 2015. Report of Sub-Committee on the Revised Management Procedure. *J. Cetacean Res. Manage.* (Suppl.) 15: 87–111.

International Whaling Commission. 2017. Report of Sub-Committee on the Revised Management Procedure. *J. Cetacean Res. Manage.* (Suppl.) 18: 123–73.

Jackson, J.A., Caroll, E.L., Smith, T.D., Zerbini, A.N., Patenaude, N.J. and Baker, C.S. 2016. An integrated approach to historical population assessment of the great whales: case of the New Zealand southern right whale. *Royal Society Open Science* 3: 150669.

Jackson, J., Olavarria, C. and Baker. 2006. Estimating the minimum historical population size of Southern Hemisphere humpback whales using diversity of mtDNA haplotypes. Paper SC/58/SH22 presented to the IWC Scientific Committee, May 2006, St. Kitts and Nevis, West Indies (unpublished). 11pp. [Paper available from the Office of this Journal].

Johnston, S.J. and Butterworth, D.S. 2010. New breeding stock C sabbatical model reference case and sensitivity tests. *J. Cetacean Res. Manage.* (Suppl. 2) 11: 244–8.

Johnston, S.J., Zerbini, A.N. and Butterworth, D.S. 2011. A Bayesian approach to assess the status of Southern Hemisphere humpback whales (*Megaptera novaeangliae*) with an application to Breeding Stock G. *J. Cetacean Res. Manage.* (Special Issue 3): 309–17.

Jones, R. 1980. Population assessments of sei whales using a modified VPA technique. *Rep. int. Whal. Commn* 30: 545–8.

King, S.L., Schick, R.S., Donovan, C., Booth, C.G., Burgman, M., Thomas, L. and Harwood, J. 2015. An interim framework for assessing the population consequences of disturbance. *Methods in Ecology and Evolution* 6(10): 1,150–8.

Kingsley, M.C.S., Richard, P. and Ferguson, S.H. 2012. Stock-dynamic model for the northern Hudson Bay narwhal population based on 1982–2008 aerial surveys. DFO Canadian Science Advisory. Secretariat Research Document. 2012/020. iv+20pp.

Kirkwood, G.P. 1981. Estimation of stock size using relative abundance data – a simulation study. *Rep. int. Whal. Commn* 31: 729–35.

Laake, J.L., Punt, A.E., Hobbs, R., Ferguson, M., Rugh, D. and Breiwick, J. 2012. Gray whale southbound migration surveys 1967–2006: an integrated re-analysis. *J. Cetacean Res. Manage.* 12: 287–306.

Leaper, R., Peel, S., Peel, D. and Gales, N. 2011. Exploring the assumptions of multi-stock assessment models for humpback whales (Megaptera novaeangliae) in the Southern Hemisphere: using Breeding Stocks D and E as an example. *J. Cetacean Res. Manage.* (Special Issue 3): 119–28.

Lennert-Cody, C.E., Rusin, J.D., Maunder, M.N., Everett, E.H., Largacha Delgado, E.D and Tomlinson, P.K. 2013. Studying small purse-seine vessel fishing behavior with tuna catch data: Implications for eastern Pacific Ocean dolphin conservation. *Mar. Mamm. Sci.* 29: 643–68.

Lewison, R.L., Crowder, L.B., Read, A. J. and Freeman, S.A. 2004. Understanding impacts of fisheries bycatch on marine megafauna. *Trends in Ecology and Evolution* 19: 598–604.

Maunder, M.N. and Punt, A.E. 2013. A review of integrated analysis in fisheries stock assessment. *Fish. Res.* 142: 61–74.

McAllister, M.K. and Ianelli, J.N. 1997. Bayesian stock assessment using catch-age data and the sampling/importance resampling algorithm. *Can. J. Fish. Aquat. Sci.* 54: 284–300.

Maunder, M.N., Starr, P.J. and Hilborn, R. 2000. A Bayesian analysis to estimate loss in squid catch due to implementation of a sea lion population management plan. *Mar. Mamm. Sci.* 16: 413–26.

Monnahan, C., Branch, T.A. and Punt, A.E. 2015. Do ship strikes threaten the recovery of endangered eastern North Pacific blue whales? *Marine Mammal Science* 31: 279–97.

Moore, J.E. and Read, A.J. 2008. A Bayesian uncertainty analysis of cetacean demography and bycatch mortality using age-at-death data. *Ecol. Appl.* 18: 1914–31.

Moore, J.E. and Barlow, J.P. 2013. Declining abundance of beaked whales (family Ziphiidae) in the California current large marine ecosystem. *PLoS ONE* 8:e52770.

Mori, M. and Butterworth, D.S. 2006. A first step towards modelling the krill-predator dynamics of the Antarctic ecosystem. *CCAMLR Sci.* 13: 217–77.

Muller, A., Butterworth, D.S. and Johnston, S.J. 2011. Assessments of the Southern Hemisphere humpback whale Breeding Stock B: results for the models and sensitivities proposed at and following the 62nd meeting of the Scientific Committee of the IWC. Paper SC/63/SH26 presented to the IWC Scientific Committee, June 2011, Tromsø, Norway (unpublished). 45pp. [Paper available from the Office of this Journal].

Nadeem, K., Moore, J.E., Zhang, Y. and Chipman, H. 2016. Integrating population dynamics models and distance sampling data: a spatial hierarchical state-space approach. *Ecology* 97: 1,735–45.

Oleson, E.M., Boggs, C.H., Forney, K.A., Hanson, M.B., Kobayashi, D.R., Taylor, B.L., Wade, P.R. and Ylitalo, G.M. 2010. Status review of Hawaiian insular false killer whales (*Pseudorca crassidens*) under the Endangered Species Act. US Dep. Commerce, NOAA Tech. Memo., NOAA-TM-NMFS-PIFSC-22. 237 pp.

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

Punt, A.E. 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(2): 127–37.

Punt, A.E. 2015a. Initial evaluation of two options for addressing infrequent surveys of the Bering-Beaufort Seas bowhead whales. Paper SC/D15/AWMP01 presented to the AWMP Workshop for Develop SLAS for the Greenland Hunts, 14–17 December 2015, Copenhagen, Denmark (unpublished). 22pp. [Paper available from the Office of this Journal].

Punt, A.E. 2015b. An investigation of the relationship between $MSYR_{MAT}$ and MSYR1+ based on a non-individual-based model. *J. Cetacean Res. Manage.* (Suppl.) 16: 110–11.

Punt, A.E. 2016. A revised age-structured model for exploring the conceptual models developed for gray whales in the North Pacific. IWC. Paper SC/A16/GW2 presented to the Third Workshop on the Rangewide Review of the Population Structure and Status of North Pacific Gray Whales, 18–20 April 2016, La Jolla, California (unpublished). 24pp. [Paper available from the Office of this Journal]

Punt, A.E. 2017. Some insights into data weighting in integrated stock assessments. *Fish. Res.* 192: 52–65.

Punt, A.E. and Breiwick, J.M. 2002. A framework for evaluating Strike Limit Algorithms for population reduced to small numbers. *J. Cetacean Res. Manage.* 4(2): 165–77.

Punt, A.E. and Butterworth, D.S. 1997. Assessments of the Bering-Chukchi-Beaufort Seas stock of bowhead whales (*Balaena mysticetus*) using maximum likelihood and Bayesian methods. *Rep. int. Whal. Commn* 47: 603–18.

Punt, A.E. and Donovan, G. 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 Polacheck, T. 2008. Further analyses related to application of statistical catch-at-age analysis to Southern Hemisphere minke whales. Paper SC/60/IA2 presented to the IWC Scientific Committee, June 2008, Santiago, Chile (unpublished). 46pp. [Paper available from the Office of this Journal].

Punt, A.E. and Wade, P.R. 2012. Population status of the eastern North Pacific stock of gray whales in 2009. *J. Cetacean Res. Manage.* 12(1): 15–28.

Punt, A.E., Butterworth, D.S., de Moor, C.L., De Oliveira, J.A.A. and Haddon, M. 2016. Management Strategy Evaluation: Best Practices. *Fish and Fisheries* 17: 303–34.

Punt, A.E., Friday, N.A. and Smith, T.D. 2006. Reconciling data on the trends and abundance of North Atlantic humpback whales within a population modelling framework. *J. Cetacean Res. Manage.* 8(2): 145–59.

Punt, A.E., Haung, T-C. and Maunder, M.N. 2013. Review of integrated size-structured models for stock assessment of hard-to-age crustacean and mollusc species. *ICES J. Mar. Sci.* 70: 16–33.

Punt, A.E., Hakamada, T., Bando, T. and Kitakado, T. 2014. Assessment of Antarctic minke whales using statistical catch-at-age analysis. *J. Cetacean Res. Manage.* 14: 93–116.

Reeves, R.R. and Brownell, R.L. (Eds). 2009. Indo-Pacific bottlenose dolphin assessment workshop report: Solomon Islands case study of *Tursiops aduncus*. Occasional Papers of the Species Survival Commission, No. 40, IUCN, Gland, Switzerland. 53pp.

Reilly, S.B. 1981. Population assessment and population dynamics of the California gray whale (*Eschrichtius robustus*). Ph.D. thesis. University of Washington, Seattle.

Reilly, S.B. 1992. Population biology and status of eastern Pacific gray whales: recent developments. pp.1,062–72. *In*: D.R. McCullough and R.H. Barrett (Eds). *Wildlife 2001: populations.* Elsevier Science, Essex, United Kingdom.

Ross-Gillespie, A., Butterworth, D.S. and Findlay, K. 2014. Assessment results for humpback breedings stocks D, E1 and oceania following recommendations from SC 65a. Paper SC/65b/SH04Rev presented to the IWC Scientific Committee, Bled, Slovenia, May 2014 (unpublished). 28pp. [Paper available from the Office of this Journal]

Ross-Gillespie, A., Butterworth, D.S. and Findlay, K. 2015. Final results for the final ’base case’ three-stock DSB, BSE1 and BSO model, with sensitivity runs. *J. Cetacean Res. Manage*. (Suppl.) 16: 215–21.

Schweder, T., Hagen, G.S. and Hatlebakk, E. 1998. On the effect on cod and herring fisheries of retuning the Revised Management Procedure for minke whaling in the greater Barents Sea. *Fish. Res.* 37: 77–95.

Shirakihara, K. and Tanaka, S. 1984. Simulation studies on length-specific population assessment techniques for western North Pacific sperm whales. *Rep. int. Whal. Commn* 34: 259–64.

Shirakihara, K., Tanaka, S. and Nakano, T. 1983. The revised age-specific model for population assessment of the Western North Pacific sperm whales. *Rep. int. Whal. Commn* 33: 757–9.

Shirakihara, K., Tanaka, S. and Nakano, T. 1985. Further simulation studies on length-specific population assessment techniques for western North Pacific sperm whales. *Rep. int. Whal. Commn* 35: 199–203.

Siler, W. 1979. A competing-risk model for animal mortality. *Ecology* 64: 750–7.

Simeoni, C. 2014. Bottlenosed dolphins in the northern Adriatic Sea: A long-term assessment of the population dynamics. MS thesis, Department of Industrial Mathematics, University of Padua. vi+103pp,

Slooten, E. 2015. Effectiveness of partial protection for Maui’s dolphin. Paper SC/66a/SM12rev1 presented to the IWC Scientific Committee, May 2015, San Diego, CA, USA (unpublished). 9pp. [Paper available from the Office of this Journal].

Smith, T.D. and Polacheck, T. 1979. Analysis of a simple model for estimating historical population sizes. *Fish. Bull.*, *US* 76: 771–9.

Smith, T.D. 1983. Changes in size of three dolphin (Stenella spp.) populations in the eastern tropical Pacific. *Fish. Bull., US* 81: 1–13.

Smith, T.D. and Reeves, R.R. 2003. Estimating historic humpback removals from the North Atlantic: an update. *J. Cetacean Res. Manage.* (Suppl.) 5: 301–11.

Stolen, M.K. and Barlow, J. 2003. A model life table for bottlenose dolphins (*Tursiops truncates*) from the Indian River lagoon system Florida, USA *Marine Mammal Science* 19: 630–49.

Tilman, M.F. 1977. Estimates of population size for the North Pacific sei whale. *Rep. int. Whal. Commn* (Special Issue 1): 98–106.

Wade, P.R. 1993. Estimation of historical population size of the eastern stock of spinner dolphin (*Stenella longirostris orientalis*). Fish. *Bull. US* 91: 775–87.

Wade, P.R. 1999. A comparison of statistical methods for fitting population models to data. pp.249–70 *In:* G.W Garner, S.C. Amstrup, J.L Laake, B.F.J. Manly, L.L. Mcdonald and D.G Robertson (Eds.) *Marine Mammal Survey and Assessment Methods*. Balkema, Rotterdam.

Wade, P.R. 2002. A Bayesian stock assessment of the eastern Pacific gray whale using abundance and harvest data from 1967–1996. *J. Cetacean Res. Manage*. 4(1): 85–98.

Wade, P.R., Reilly, S.B. and Gerrodette, T. 2002. Assessment of the population dynamics of the northeastern offshore spotted dolphin and the eastern spinner dolphin populations through 2002. SWFSC Administrative Report LJ-02-13. 58pp.

Wade, P.R., Watters, G.M. Gerrodette, T. and Reilly, S.B. 2007. Depletion of spotted and spinner dolphins in the eastern tropical Pacific: modeling hypotheses for their lack of recovery. *Mar. Ecol. Prog. Ser.* 343: 1–14.

Witting, L. 2013. Assessment runs for harbour porpoise in West Greenland. NAMMCO Document. NAMMCO/SC/20/HP/05. 24pp.

Witting, L. 2015. Meta population modelling of narwhals in East Canada and West Greenland. NAMMCO Document NAMMCO/SC/22/JCNB/SWG/2015-JWG/10. 20pp.

Witting, L. and Heidi-Jørgensen, M.P. 2015a. Population model for West Greenland beluga. NAMMCO Document. NAMMCO/SC/22/JCNB/SWG/2015-JWG/09. 12pp.

Witting, L. and Heidi-Jørgensen, M.P. 2015b. Assessment of East Greenland narwhals. NAMMCO Document NAMMCO/SC/22/JCNB/SWG/2015-JWG/16. 14pp.

Zerbini, A.N., Clapham, P.J. and Wade, P.R. 2010. Assessing plausible rates of population growth in humpback whales from life-history data. *Mar. Biol*. 157: 1,225–36.

Zerbini, A.N., Ward, E.J., Kinas, P.G., Engel, M.H. and Andriolo, A. 2011. A Bayesian assessment of the conservation status of humpback whales (*Megaptera novaenangliae*) in the western South Atlantic Ocean (Breeding Stock A). *J. Cetacean Res. Manage.* (Special Issue 3): 131–44.

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(2): 169–75.