# Review of Contemporary Cetacean Stock Assessment Models

**André E. Punt**
**School of Aquatic and Fishery Sciences, University of Washington, Seattle, WA 98195**

## Workshop on Methods for Monitoring the Status of Eastern Tropical Pacific Ocean Dolphin Populations
## 18-20 October 2016, La Jolla, California

# DRAFT

# 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 assessments, and to provide the operating models on which management strategy evaluation is based, for cetacean stocks. 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. Abundance is very low for some stocks, so both demographic and environmental variability need to be included in models for these stocks. The primary source of data for parameter estimation is time-series of estimates of absolute abundance, although the analyses reviewed made use 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 because there is a lack of catch-induced declines in abundance that are captured by indices of relative abundance and hence 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. However, some analyses also quantify uncertainty using bootstrap and asymptotic methods. The future for the models on which assessments and management strategy evaluation is based will likely 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.

## 1. 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 from those approaches applied to fish and invertebrates in some significant ways. Specifically, catches (at least during the most recent three decades) have tended to be low for most cetacean stocks – generally only bycatch, and in a few instances commercial and aboriginal 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.

The assessments of cetacean populations are used for a variety of purposes. Specifically, they can be used to provide (a) information on abundance in absolute terms and relative to the pre-exploitation size and to target and threshold levels, (b) estimates of recent trends in abundance and/or mortality, and (c) probabilities of rebuilding and extinction. Management advice for several cetacean stocks are based on the application of management strategies[2]. In a few cases (e.g., for dolphin stocks off the North American west coast) the outputs from the assessments provide the estimates of abundance that are used to calculate catch limits. However, in most of the cases where catch limits (or strike limits) are set for cetacean stocks, these are based on management strategies that use survey-based estimates of abundance, empirical rules that 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 is usually based on simulation testing; a core element of simulation testing is the population dynamics model that represents the truth for the simulations (i.e., 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" are 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 next section of this paper lists all of the stocks for which analyses have been undertaken and the analysis methods used most recently for those analyses. The focus is on analysis methods rather than the results of the analyses or even whether the 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).

## 2. Stocks and analyses

The review focuses on recent (generally since 1995) analysis methods that involve population dynamics models that were applied to cetacean stocks. Thus, it does not cover the models

---

[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.

[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). Often referred to as "management procedures" in the cetacean literature.

used to analyse the monitoring data used to provide the estimates of abundance on which stock assessments are based (e.g., Gerrodette and Forcada, 2005; Canadas *et al*., 2006), the models used to standardize catch-per-unit effort data (e.g., Cooke, 1993), and the models used to analyse mark-recapture data[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 / guestimated (e.g., Alvarez-Flores, 2006; Dueck and Richard, 2008; Reeves and Brownell, 2009; Slooten, 2015) are not covered in this review. Similarly, models that are only approximately fitted to data and were developed primarily to estimate life history parameters (e.g., Fifas *et al*., 1998; Sloten and Barlow, 2003) are not considered in this review.

The set of stocks, and hence the analysis methods, summarized in this review were identified through a literature search (Web of Science / Google Scholar), contacts with representatives of key management bodies, as well as contacts with individual analysts. Many of the reports describing analyses are found in the gray literature so are not necessarily searchable in databases such as web of science.

The results for baleen and sperm whales 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 Scientific Committee of the International Whaling Commission, 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 North Atlantic Marine Mammal Commission, NAMMCO) review processes. The information collected is summarized by ocean basin or by stock, depending on the unit of analysis. In some cases, a stock has been assessed as a single unit and as part of a regional analysis. In these cases, results are presented separately for the single unit and regional analyses.

Tables 1 and 2 lists the stocks / species considered in this review, their major purpose (to form an assessment or to be the operating model for an MSE), the basic structure of the model, and some key references. The key references tend to be 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 Southern Hemisphere 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 summarize 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.

## 3. Model structure assumptions

### *3.1 Population dynamics models*

The assessments in Tables 1 and 2 are based on several types of population dynamics models. 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 spinner dolphins, and eastern spotted dolphins). These analyses provide no information about the status of stocks relative to reference points such as carrying capacity (except perhaps whether populations are increasing or not).

Most of the analyses in Table 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

---

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

production occurs (MSYL = 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 Maunder (2004) represented the population of eastern 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 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 very computationally intensive, especially if the aim is to quantify uncertainty using bootstrap and/or Bayesian methods so their application has to date been limited to small populations (<1,000 animals in total) for which resighting probabilities are at least 10%.

Most of the analyses are for a single stock and 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 whales, gray whales, minke whales, and humpback whales to account for catches on feeding grounds likely consisting of multiple stocks, and there being no objective way to assign catches on, and estimates of abundance for, feeding grounds to stocks. Other reasons 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 model structure 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 *et al*., 1983; 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 (such as struck and lost rates are markedly in error). 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) did not make the assumption that carrying capacity has been constant for 150 years and instead, following Wade (2002), started the population projections in 1940, with the age-structure at that time assumed to be stable. Punt and Butterworth (2002) started population projections 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 whalers seldom appear to select for animals of particular ages / sizes (and catch data are often available by sex anyway). However, spatial variation in age structure may interact with the spatial distribution of the fisheries to produce apparent spatial and temporal differences in selectivity. There are some analyses with multiple fleets. Examples include the analyses for the eastern North Pacific stock of gray whales and minke whales off West Greenland, which include multiple fleets owing to differences in selectivity patterns between commercial and aboriginal whalers. Multiple fleets are 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 minke whales in the southern hemisphere. The latter assessment allows for time-varying commercial selectivity given among-year changes in where the various fisheries operated.

### *3.2 Density-dependence*

Density-dependence could operate on a variety of population processes. For example, density-dependence could impact maturation, growth, calving rate, juvenile survival, adult survival and perhaps 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) compare several alternative 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) 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.

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

$$C_y = N_y^m \, f_0 (1 + A(1 - (N_y^d / K^d)^z) \qquad (1)$$

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 (MSYR, the ratio of MSY to the equilibrium number of recruited animals when the population is

---

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

producing MSY). 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 carrying capacity, 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 minimization is based on software that requires a differentiable objective function (such as AD Model Builder, Fournier *et al*. [2012]). Consequently, the assessment of Southern Hemisphere 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) or non-calf survival (or both) (Punt, 2015b). However, only one of the analyses (that for Cook Inlet Beluga whales, Hobbs and Sheldon, 2008; Hobbs *et al*., 2016) included density-dependent natural mortality.

### *3.3 Other population dynamic assumptions*

The base versions of the analyses are generally quite similar, but there are often many differences in the alternative models examined to conduct tests of sensitivity. The focus here is on the assumptions for the base versions of the models. Key differences among the models include:

- *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 minke whales in the Southern Hemisphere, which allows for deviations in recruitment about the density-dependence function (i.e., about expected calf numbers), in the proportion of the population in each area in which the two stocks of minke whales are found, in deviations in selectivity spatially and over time, and in carrying capacity. Several other assessments (generally of shorter-lived species) consider stochastic recruitment, including the model developed Hoyle and Maunder (2004) for eastern spotted dolphins, that for Cook Inlet Beluga whales, and that for Hectors 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. 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 is pre-specified (Tables 5 and 6), in some cases, natural mortality depends on age (e.g., for fin and minke whales in the North Atlantic and North Pacific). Some of the analyses estimate natural mortality (and in the case of Southern Hemisphere 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. 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.
- *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 very 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 carrying capacity 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 carrying capacity 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 carrying capacity 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 carrying capacity occurred (for unknown reasons). The assessment of Southern Hemisphere minke whales estimates changes over time in carrying capacity as a random walk, thereby avoiding having to specify (or estimate) when carrying capacity changed. Estimation of MSYR is challenging even when it is assumed to be time-invariant. Consequently, consideration of time-varying productivity is unusual[5]. However, the analyses of dolphin populations in the eastern tropical Pacific considered model variants that estimated two levels for MSYR (modelled as the intrinsic rate of growth), i.e. implicitly assuming that a regime in productivity occurred.
- *How is selectivity modelled?* The choice of the fishery selectivity pattern is likely 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 North Atlantic minke whales). However, the availability of age-composition data has allowed selectivity to be estimated for some stocks (Southern Hemisphere minke whales, North Atlantic fin whales, the Bering-Chukchi-Beaufort Seas stock of bowhead whales [Punt, 2006], sperm whales in the western North Pacific, spotted dolphins in the eastern tropical Pacific, and narwhals and harbor porpoise off West Greenland). The assessment of minke whales in the 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 likely 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.

---

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

- *How is the assessment 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.

The models that consider spatial structure almost always do not represent spatial structure explicitly, i.e. no attempt is made to define the probability that whales in one area move to another areas. Rather, the models that consider spatial structure estimate (or pre-specify) the proportion of each stock in each area, with the estimates of the mixing proportions based primarily on data on the proportion of each stock in each area from, for example, genetics information. In general, the models that include multiple stocks assume that there is no permanent transfer of animals between stocks ("diffusion"). Exceptions to this general rule are the models developed to test management strategies for 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 are based on models with an annual time-step. The exception is the model on which the MSE for the western North Pacific minke whales is based, which operated on a monthly time-step to capture the impact of harvesting during a migration.

## 4. 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 assessments (IWC, 1989).

Catches were included in most of the analyses (Tables 3 and 4). However, catches, particularly those for the earliest years of exploitation, 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 commercial and aboriginal harvesting, although the model used for rangewide assessment of Pacific gray whales by 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 the situation for baleen and sperm whales, the bulk of the anthropogenic removals of dolphins are due to bycatch. Bycatch estimates are usually much more uncertain that catches by commercial whaling (e.g., Lo and Smith, 1986).

All but one of the analyses made use of estimates of absolute abundance for parameter estimation purposes. A noteworthy exception was the models developed for sperm whales in the western North Pacific, which were fitted to the catch length-frequency for males. Those models were developed in the early 1980s, prior to the start of most of the major survey programs. Consequently, were the assessments of western North Pacific sperm whales to be revisited, they would likely use survey estimates of abundance (perhaps as relative indices of abundance given difficulties estimating g(0) for 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 should not be the case when sample sizes are small so 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. Moore and Barlow (2013) 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 catchability for the surveys, often because the g(0) 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). Consequently, those assessments selected the value for carrying capacity so that model “hit” the available estimate abundance (de la Mare, 1989). However, as additional surveys were conducted, it was possible to include the abundance data in the likelihood function maximized 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 animals. This has led to the practice of estimating an “additional variance” parameter for surveys. Additional variance is now commonly estimated in analyses in which there are multiple estimates of absolute or relative abundance. Such additional variation may represent sampling error, temporal variation in survey catchability, unmodeled 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 needs to 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 mixing rates for North Atlantic fin whales and western North Pacific Bryde’s whales, to estimate abundance for southwest and southeast Atlantic right whales, gray whales off Sakhalin Island, and several of the stocks of humpback whales in the Southern Hemisphere, and to estimate survival for Hector’s dolphins off Bank’s Peninsula. In principle, mark-recapture data can be used to estimate abundance. However, 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 the mark-recapture data may imply a non-significant trend in abundance, but a statistically significant trend may be detected if these data 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 humpback whale stocks in the Southern Hemisphere included a constraint on the lower bound for the total number of animals in the population based on counts of mtDNA haplotypes. As noted by Jackson *et al*. (2006), the observed number of haplotypes in a population provides an absolute minimum on the number of females when the population was at its lowest level. 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 the population being at its lowest level. In general, the impact of imposing an $N_{\text{min}}$ is greatest when it is large because $N_{\text{min}}$ 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 age- and size-composition 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). Care needs to be taken when including age- and length-composition data in analyses because these data can provide information on absolute abundance, but the information is very 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).

## 5. 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.

### *5.1 Measures of statistical uncertainty*

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

The bulk 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, including that some of the first uses of Bayesian methods to conduct assessments of marine populations subject to harvest occurred for cetaceans (e.g., Givens *et al*., 1995) so there is a historical precedent for the use of Bayesian methods for this group of species, and that production of posterior distributions is computationally feasible for many cetacean stocks given the relatively limited amount of data for most such stocks. More importantly perhaps is that Bayesian methods provide a way to include prior information in analyses, particularly because of the limited amount of information contained in the data for most stocks (e.g. for the MSYL). Priors can be assumed to be uniform (e.g., Wade *et al*. 2002, 2007). However, it is preferable to base a Bayesian analysis on priors that are informative and represent a synthesis of parameter estimates among species and stocks (i.e., the analysis is based on “data-based” priors). Most of the analyses in Tables 1 and 2 based on Bayesian methods imposed priors on biological parameters such as the age-at-maturity, the 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 equivalently the maximum growth rate). However, in many cases, there is little information to update the priors (e.g., the eastern North Pacific blue whales), and in some cases, priors are updated to values that are biologically unrealistic or implausible. 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 by the Scientific Committee of the IWC. 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 is well known. In the context of assessments of cetaceans, the key discussions have related to whether it is reasonable to impose independent priors on each of the biological parameters age-at-maturity, survival rate and maximum pregnancy rate given observed correlations between the values for the parameters when estimates can be made, 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 should a prior be imposed on carrying capacity or abundance in a recent year[6]. In general, while data can update the prior for carrying capacity (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 catches tend to be low compared to productivity in most cases, particularly during recent years when most of the monitoring data are available. Therefore, information on absolute abundance contained in catch-induced changes in relative abundance is not available. However, parameters related 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 (c.f., IWC, 2015; Tables 3 and 4).

### *5.2 Sensitivity analysis*

All but one of the analyses examine sensitivity to assumptions using sensitivity analyses in which some of the assumptions of a base model (or a set of base models) are changed. The exploration of sensitivity tends to be most extensive for the management strategy evaluations because one objective of MSE is to identify a management strategy that is robust to the uncertainty. The aim when designing an MSE is that the set of operating models will be reduced and not increased 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. However, it is seldom the case that even MSEs will explore *all* plausible hypotheses and assumptions. Nevertheless, the number of sensitivity tests can be

---

[6] 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.

substantial for some MSEs (see Table 9 for the sensitivity tests conducted for the MSE for the Bering-Chukchi-Beaufort Seas stock of bowhead whales). The sensitivity tests for MSEs in which there is uncertainty regarding stock structure can involve changing the number of stocks in the region being managed and where they are located (e.g., fin and minke whales in the North Atlantic and 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.

### *5.3 Simulation evaluation*

It is now best practice in resource management to evaluate the performance of assessment methods before they are used to provide management advice. The Scientific Committee of the International Whaling Commission 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 whales 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 very computationally extensive. However, there are some examples of recent assessment methods (including Bayesian methods) having been evaluated using (often limited) simulation: (a) the Bering-Chukchi-Beaufort Seas stock of bowhead whales (Punt and Butterworth, 1997), (b) 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).

## 6. Projections and management outputs

Most, but not all, of the analyses have the capability to conduct projections (Tables 7 and 8). The models developed as the basis for operating models to evaluate alternative management strategies 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 carrying capacity. The most extensive evaluation of the future state of a cetacean population was conducted by Hobbs *et al*. (2016) for beluga 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. However, they did not estimate posterior distributions for all of these processes, but instead examined sensitivity to alternative plausible values for the parameters governing them. The assessment of Southern Hemisphere minke whales reported time-trends in calf numbers, as well as growth rates and carrying capacity. This information is not reported for other assessments because they do not estimate changes over time in recruitment, growth and carrying capacity.

In contrast to the assessments, the MSEs evaluate full-feedback management strategies. Thus, the MSEs include a component that generates the types of data that will be available in the future to form the basis for assessments. In general, these are estimates of absolute abundance, but could include other information such as the proportion of the population that are calves, juveniles or adults (e.g., the MSE developed for the Bering-Chukchi-Beaufort Seas stock of bowhead whales; IWC, 2003). The relative lack of data generated as part of the

MSEs is in contrast with the MSEs developed to evaluate management strategies for fisheries management 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. 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 Revised Management Procedure used to specify catch limits for baleen whales that are caught on their feeding grounds by commercial whalers, IWC, 2012) 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 MSE for bowhead whales off West Greenland where account is taken of catches by Canada. The evaluation of variants of the Revised Management Procedure for fin and minke whales off west and east Greenland was based on MSEs that pre-specified the catches in aboriginal hunts.

The common outputs from analyses (and their projections) are time-trajectories of numbers of animals in absolute terms or relative to carrying capacity (or other reference points such as MSYL). The population numbers are usually summarized as the total population size, although some assessments also report numbers of females (e.g., Cooke *et al*., 2016) or even 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., Given *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 *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.

## 6. Discussion

### *7.1 Best practices for modelling cetacean stocks*

Table 10 lists a set of "best practice" guidelines for conducting analyses for cetacean stocks

#### 7.1.1 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 Beluga whales and Banks Peninsula Hector's dolphins) as well as to populations consisting of thousands to hundreds of thousands of individuals (e.g., minke whales in Southern Hemisphere, 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. Unlike fish and invertebrates, the number of calves-per-female is constrained for a cetacean. Consequently, there are limits to the amount by which the number of calves can differ from the expected value given by equations such as equation 1. Punt *et al*. (2014) recognized this, and formulated the function defining recruitment variation to impose an upper bound on the numbers of calves-per-female in any year. In general, stochasticity in calf numbers has limited impact on population trajectories when calf survival is larger than 0.9. However, this type of 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 variables 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 computational, 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.

There is often little justification for the inclusion of sex-structure in analyses. However, it would be prudent to explicitly model sex-structure for species for which the catch sex-ratio can be markedly different from 1:1 (such as minke 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 mature female abundance, but social behaviour related to reproduction might result in the number of males impacting reproduction rates.

Most of the early analyses assumed that the region under consideration contained only a single stock. However, 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. None of the models that allow for multiple areas and movement, model movement explicitly. Rather these models treat the proportion of each stock in each modelled areas as estimable parameters (or pre-specify these parameters).

Finally, most models ignore within-year dynamics. This is generally 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 stock of minke whales was the only one that allowed for seasonal dynamics. This structure was needed because catches occur during migration, and consequently the stock-, sex-, and age-composition of the catches in some areas changes during the season.

In general, estimation performance, measured by the precision with which parameters such as carrying capacity is estimated, is improved if the stock is assumed to be at carrying capacity 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.

### 7.1.2 Parameterization 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.

Clearly, 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 this should remain the default for analyses. However, the impacts of density-dependence in adult survival are such that this source of density-dependence is worth at least considering in analyses.

Natural mortality is likely age-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). 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 with breakpoints based on the results of other models for natural mortality-at-age.

How selectivity is modelled is generally likely to be inconsequential owing to the longevity of most cetaceans. However, selectivity should be estimated rather than being pre-specified if historical removals were very large and particularly if age- (or size-) composition data are included in the likelihood. 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, it is worth testing this assumption, especially 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, carrying capacity and distribution could be linked (perhaps with error - Brandon and Punt [2013] - to environmental variables). 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 an environmental variable.

### 7.1.3. 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 management strategy evaluations are based. The uncertainties that usually have the greatest impact on estimates of current abundance, and current abundance relative to reference points are: (b) 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 catchability coefficient for estimates of abundance, 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 presented is plausible. This is particularly the case for uncertainties that relate to possible future events (e.g., future changes in carrying capacity and productivity, and 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, it is advisable to divide sensitivity tests into a reference set that consists of the most likely sets of assumptions and a robustness set that includes scenarios that are of interest, but are not very likely. This approach has been taken by the Scientific Committee of the International Whaling Commission for several recent MSEs.

A variety of ways exist to quantify uncertainty (Tables 7 and 8). However, 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).

### 7.1.4. Data

In general, it is better to use as many sources of data as possible in assessments. However, model misspecification, including incorrect assumptions about sampling error, can degrade results when multiple data sources are used for parameter estimation. Another exception to this general recommendation is that CPUE data are unlikely to be representative of changes in population size. Inclusion of multiple data sources, can, however, 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 should be available 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).

The availability of multiple sources of data raises the possibility that those data sources are in conflict to some extent. This leads to the need to weight the data sources (or completely ignore some of them). Data weighting is a core component of stock assessment in fisheries and needs to be objective and replicable. Some of the methods used when conducting fisheries stock assessments have been used in the assessments that have used, for example, included both index and age- and length-composition data (e.g. Punt *et al*. 2014).

### *7.3 Key future directions*

The analyses outlined in the paper ignore biological interactions among species. Such interactions could occur due to whales feeding on a common prey base. Mori and Butterworth (2006) outline a modelling framework based on a biomass dynamics model for Southern Hemisphere blue, fin, humpback and minke whales feeding on krill. That model also includes two seal species (Antarctic fur and crabeater seals). In common with the many multi-species models, it was difficult to fit all of the available data simultaneously, and this analysis has not formed the basis for management advice. Schweder *et al*. (1998) also developed a multi-species model involving cetaceans and pinnipeds, but in common with the work of Mori and Butterworth (2006), it has not formed the basis for management advice.

The models on which analyses for cetacean stocks are based include many of those conventionally used as the basis for fisheries stock assessments. In principle, given the population sizes of some of the stocks involved, it would be possible to apply individual-based models. This is essentially how the mark-recapture-based assessments for southwest and southeast Atlantic right whales and Sakhalin Island gray whales are formulated. Punt and

Breiwick (2002) outline 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.

Baker and Clapham (2004) indicated that a key future trend for assessment of cetaceans was likely to be the increased use of genetic data for population estimation. However, while genetic data are being used to provide minimum estimates of population size, form the basis for genetic mark-recapture studies, and inform mixing rates when multiple stocks of a species are found in a single area, the promise of using diversity data to estimate population size has not be realized yet.

Increasingly data sources are being included in assessments in their raw form. Examples of this are the models used for right whales in the southwest and southeast Atlantic and for gray whales off Sakhalin Islands that integrate mark-recapture histories directly into the population model. However, most assessments fit the population model to estimates of abundance when these are determined from surveys 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 in Nadeem *et al*. (2016) is based on an age- and sex-aggregated model, with production based on a Gompertz model and no allowance for historical removals. In principle, the approach of Nadeem *et al*. (2016) could be extended to account for age, sex and catches but this might come at a substantial computational cost.

Finally, 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. This perhaps reflects the complexity of some of the estimation methods. However, it is counter to the trend in fisheries assessment where most of the key methods have been subject to some form of simulation evaluation (Dichmont *et al*., 2016).

### *7.4 Final thoughts*

Model-based assessments of cetaceans remain the gold standard for providing management advice. In comparison with assessments for fish and invertebrates, 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 is inferred from changes in relative abundance and age-composition. However, estimation of trends in abundance (and hence the values for parameters such as MSYR) rely on information such as trends in relative abundance or age-composition, which are often unavailable 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 very 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 programs. 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.

The analyses for cetaceans are similar in some respects with those used in fisheries and it might be expected that this would also be the case for analyses for pinnipeds. This is true to some extent, with analyses for pinnipeds often based on models that include density-dependence in pup survival (e.g. Butterworth *et al*., 1995; Breen *et al*., 2003; Fay and Punt, 2006; Towell, 2007), although Towell (2007) allows density-dependence to impact the survival rates of several of the younger age-classes. These models tend to be based on age- and sex-structured models (but see Maunder *et al.*, 2000, for an exception in this regard). The models for pinnipeds tend to be fitted to indices of pup numbers (as this class is easiest to survey), but they can be fitted to many of the same data sources used in assessments of cetaceans, fish and invertebrates such as catch age-composition. However, they make use of data sources that are somewhat unique to these species such as estimates of survival (Fay and Punt, 2006), counts of adults (Towell, 2007). In addition, there is a much greater focus on pup dynamics in pinniped models (see, for example, Butterworth *et al*., 1995), which is not a well-monitored class for many cetaceans (although there are exceptions such as right and gray whales).

Use of MSE to evaluate candidate management strategies is becoming standard in fisheries science and is central to the work of, for example, the Scientific Committee of the IWC. However, MSEs need to be tailored to the situation in question. For example, the evaluation of Strike Limit Algorithms for aboriginal whaling focus on whether the resulting removals will allow the stock to (continue to) rebuild rather the probability that it is some function of carrying capacity or MSYL, owing to the general inability to estimate carrying capacity, and hence quantities that are functions of carrying capacity such as MSYL, for the stocks concerned.

This review has contrasted cetacean analyses with those for fish and invertebrates, noting that there are generally fewer data available for parameter estimation purposes for cetaceans. However, fisheries science has much to learn from analyses conducted for cetaceans, in particular the way MSE has been applied, the non-use of commercial catch rate information, and by the attempts to better understand the implications of alternative stock structure hypotheses.

## 8. Acknowledgements

This work was conducted under contract to the Inter-American Tropical Tuna Commission. Mark Maunder, Cleridy Lennert (IATTC) and Malcolm Haddon (CSIRO and Oceans and Atmosphere) are thanked for comments on early version of this paper.

## 9. 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. Secretariat 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 International Dolphin Conservation Program. 12pp.

Baker, C.S. and P.J. Clapham. 2004. Modelling the past and future of whales and whaling. *Trends in Ecology and Evolution* 19: 365-371.

Beddington, J.R. and J.G. Cooke. 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. *Reports of the International Whaling Commission* 31: 747-760.

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. *Journal of Cetacean Research and Management* (*Supplement*) 12: 107-110.

Brandao, A., Butterworth, D.S., Ross-Gillespie, A. and P.B. Best. 2013. Application of a photo-identification based assessment model to southern right whales in South African waters, now including data up to 2012. IWC Document SC/65/BRG17. 15pp.

Brandon, J.R. and A.E. Punt. 2013. Testing the *Gray Whale* Strike Limit Algorithm (*SLA*): allowing environmental variability to influence population dynamics. *Journal of Cetacean Research and Management* 13: 81-88.

Brandon, J.R. and P.R. Wade. 2006. Assessment of the Bering-Chukchi-Beaufort Seas stock of bowhead whales using Bayesian model averaging. *Journal of Cetacean Research and Management* 8: 225-239.

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

Breen, P.A., Hilborn, R., Maunder, M.N. and S.W. Kim. 2003. Effects of alternative control rules on the conflict between a fishery and a threatened sea lion (*Phocarctos hookeri). Canadian Journal of Fisheries and Aquatic Sciences* 60: 527–541.

Breiwick, J.M. 1977. Analysis of the Antactic fin whale stock in area I. *Reports of the International Whaling Commission* 27: 124-127.

Butterworth, D.S. and A.E. Punt. 1992. Assessments of the East Greenland-Iceland fin whale stock. *Reports of the International Whaling Commission* 42: 671-696.

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

Butterworth, D.S., Punt, A.E., Oosthuizen, W.H. and P.A. Wickens. 1995. The effects of future consumption by the Cape fur seal on catches and catch rates of the Cape hakes. 3. Modelling the dynamics of the Cape fur seal population. *South African Journal of Marine Science* 16: 161-183.

Butterworth, D.S., Punt, A.E. and A.D.M. Smith. 1996. On plausible hypotheses and their weighting, with implications for selection between variants of the Revised Management Procedure. *Reports of the International Whaling Commission* 46: 637-640.

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

Canadas, A., Fortuna, C.M. and P.S. Hammond. 2006. Modelling techniques to investigate the impact of changes in habitat on cetacean distribution and abundance. IWC Document SC/58/E21. 11pp.

Cooke, J.G. 1986. On the net recruitment rate of gray whales with reference to inter-specific comparisons. *Reports of the International Whaling Commission* 36: 363-366.

Cooke, J.G. 1993. Further analyses of minke whale CPUE in the northeast Atlantic with estimate of yield rates. IWC Document SC.45/NA7. 18pp.

Cooke, J.G. and W.K. de la Mare. 1983. Description of and simulation studies on the length-specific sperm whale assessment technique. *Reports of the International Whaling Commission* 33: 741-746.

Cooke, J., Rowntree, V. and R. Payne. 2003. Analysis of inter-annual variation in reproductive success of South Atlantic right whales (*Eubalaena australis*) from photo-identifications of calving females observed off Peninsula Valdes, Argentina, during 1971-2000. IWC document SC/55/O23.16pp.

Cooke, J.G., de la Mare, W.K. and J.R. Beddington. 1983. An extension of the sperm whale model for the simulation of the male population by length and age. *Reports of the International Whaling Commission* 33: 731-734.

Cooke, J.G., Weller, D.W., Bradford, A.L., Sychenk, O., Burdin, A.M., Lang, A.R. and R.L. Brownell, Jr. 2016. Updated Population Assessment of the Sakhalin Gray Whale Aggregation based on a photoidentification study at Piltun, Sakhalin, 1995-2015. IWC Document SC/66b/BRG225. 13pp.

Davies, N., Bian, R., Starr, P., Lallemand, P., Gilbert, D. and J. Mckenzie. 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. 113pp.

de la Mare, W.K. 1986. Fitting population models to time series of abundance data. *Reports of the International Whaling Commission* 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. *Reports of the International Whaling Commission* 38: 243-248.

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). *Reports of the International Whaling Commission* 39: 150–151.

de al Mare, W.K. 2016. Simulation studies on the properties of estimates form statistical catch at age models. IWC Document SC/66b/IA8. i+34pp.

de la Mare, W.K. and J.G. Cooke. 1984. A new model for the female sperm whale incorporating aspects of the reproductive cycle. *Reports of the International Whaling Commission* 34: 251-254.

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 K. Shertzer. 2016. A review of stock assessment packages in the United States. *Fisheries Research* 183: 447-460.

Dueck, L. and P. Richard. 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.

Fay, G. and A.E. Punt. 2006. Modeling spatial dynamics of Steller sea lions (*Eumetopias jubatus*) using maximum likelihood and Bayesian methods: evaluating causes for population decline. p. 405-433. In: A.W. Trites, S.K. Atkinson, D.P. DeMaster, L.W. Fritz, T.S. Gelatt, L.D. Rea, and K. Wynne [Ed.]. *Sea lions of the world*. Alaska Sea Grant College Program, University of Alaska Fairbanks.

Fifas, S., Goujon, M. and L. Antoine. 1998. Application of Leslie's model on a population of common dolphins (*Delphinus delphis*): Senstivity study. *Aquatic Living Resources* 11: 359-369.

Fournier, D.A., Skaug, H.G., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M.N., Nielsen, A., and J. Sibert. 2012) AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. *Optimization Methods and Software* 27: 233–249.

Francis, R.I.C.C. 2011. Data weighting in statistical fisheries stock assessment models. *Canadian Journal of Fisheries and Aquatic Sciences* 68: 1124–1138.

Gerrodette, T. and J. Forcada. 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 A.E. Raftery. 1995. Assessment of the Bering Chukchi-Beaufort Seas stock of bowhead whales using the BALEEN II model in a Bayesian Synthesis Framework. *Reports of the International Whaling Commission* 45: 345-364.

Government of Japan. 2016. Results of the analytical work on NEWREP-A recommendations on sample size and relevance of age information for the RMP. IWC Document SC/66b/SP10. 23pp.

Hammill, M.O., Kingsley, M.C.S, Lesage, V. and J-F. Gosselin. 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. *Reports of the International Whaling Commission* 30: 519-530.

Hobbs, R.C. and K.E.W. Shelden. 2008. Supplemental status review and extinction assessment of Cook Inlet belugas (*Delphinapterus leucas*). AFSC Processed Rep. 2008-08, 7Alaska Fish. Sci. Cent., NOAA, Natl. Mar. Fish. Serv., 7600 Sand Point Way NE, Seattle WA 98115. 76pp.

Hobbs, R.C, Wade, P.R. and K.W. Sheldon. 2016. Viability of a Small, Geographically-isolated Population of Beluga Whales, *Delphinapterus leucas*: Effects of Hunting, Predation, and Mortality Events in Cook Inlet, Alaska. *Marine Fisheries Review* 77: 59-88.

Hoyle, S.D. and M.N. Maunder. 2004. A Bayesian integrated population dynamics model to analyse data for protected species. *Animals Biodiversity and Conservation* 27: 247-266.

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

International Whaling Commission. 2003. Report of Standing Working Group on the Aboriginal Subsistence Whaling Management Procedure. Annex E to Report of the Scientific Committee. *Journal of Cetacean Research and Management* (*Supplement*) 5: 154-225.

International Whaling Commission. 2005. Report of Standing Working Group on the Aboriginal Subsistence Whaling Management Procedure. Annex E to Report of the Scientific Committee. *Journal of Cetacean Research and Management* (*Supplement*) 7: 115-187.

International Whaling Commission. 2008a. Report of Standing Working Group on the Aboriginal Subsistence Whaling Management Procedure. Annex E to Report of the Scientific Committee. *Journal of Cetacean Research and Management* (*Supplement*) 8: 121-149.

International Whaling Commission. 2008b. Report of Sub-Committee on the Revised Management Procedure. Annex D to Report of the Scientific Committee. *Journal of Cetacean Research and Management* (*Supplement*) 10: 90-119.

International Whaling Commission. 2012. The Revised Management Procedure (RMP) for Baleen Whales. *Journal of Cetacean Research and Management* (*Supplement*) 13: 485–494.

International Whaling Commission. 2013. Report of Standing Working Group on the Aboriginal Subsistence Whaling Management Procedure. Annex E to Report of the Scientific Committee. *Journal of Cetacean Research and Management* (*Supplement*) 13: 130-153.

International Whaling Commission. 2014a. Report of Standing Working Group on the Aboriginal Subsistence Whaling Management Procedure. Annex E to Report of the Scientific Committee. *Journal of Cetacean Research and Management* (*Supplement*) 15: 189-213.

International Whaling Commission. 2014b. Report of Sub-Committee on the Revised Management Procedure. Annex D to Report of the Scientific Committee. *Journal of Cetacean Research and Management* (*Supplement*) 15: 87-188.

International Whaling Commission. 2015. Report of Sub-Committee on the Revised Management Procedure. Annex D to Report of the Scientific Committee. *Journal of Cetacean Research and Management* (*Supplement*) 15: 87-111.

International Whaling Commission. 2017. Report of Sub-Committee on the Revised Management Procedure. Annex D to Report of the Scientific Committee. *Journal of Cetacean Research and Management* (*Supplement*) 18: 00-00.

Jackson, J.A., Caroll, E.L., Smith, T.D., Zerbini, A.N., Patenaude, N.J. and C.S. Baker. 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.A., Olavarria, C. and C.S. Baker. 2006. Estimating the minimum historical population size of southern hemisphere humpback whales using diversity of MtDNA haplotypes. IWC Document SC/58/SH22. 11pp.

Johnston, S.J. and D.S. Butterworth. New breeding stock C sabbatical model reference case and sensitivity tests. *Journal of Cetacean Research and Management* (*Supplement 2*) 11: 244-248.

Johnston, S.J., Zerbini, A.N. and D.S. Butterworth. 2011. A Bayesian approach to assess the status of Southern Hemisphere humpback whales (*Megaptera novaeangliae*) with an application to Breeding Stock G. *Journal of Cetacean Research and Management* (Special Issue) 3: 309-317.

Jones, R. 1980. Population assessments of sei whales using a modified VPA technique. *Reports of the International Whaling Commission* 30: 545-548.

Kingsley, M.C.S., Richard, P. and S.H. Ferguson. 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. *Reports of the International Whaling Commission* 31: 729-735.

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

Leaper, R., Peel, S., Peel, D. and N. Gales. 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. *Journal of Cetacean Research and Management* (Special Issue) 3: 119-128.

Lennert-Cody, C.E., Rusin, J.D., Maunder, M.N., Everett, E.H., Largacha Delgado, E.D and P.K. Tomlinson. 2013. Studying small purse-seine vessel fishing behavior with tuna catch data: Implications for eastern Pacific Ocean dolphin conservation. *Marine Mammal Science* 29: 643-668.

Lo, N.C.H. and T.D. Smith. 1986. Incidental mortality of dolphins in the eastern tropical Pacific, 1959-72. *Fishery Bulletin* 84: 27-34.

Maunder, M.N. and A.E. Punt. 2013. A review of integrated analysis in fisheries stock assessment. *Fisheries Research* 142: 61-74.

McAllister, M.K. and J.N. Ianelli. 1997. Bayesian stock assessment using catch-age data and the sampling/importance resampling algorithm. *Canadian Journal of Fisheries and Aquatic Sciences* 54: 284–300.

Maunder, M.N., Starr, P.J. and R. Hilborn. 2000. A Bayesian analysis to estimate loss in squid catch due to implementation of a sea lion population management plan. *Marine Mammal Science* 16: 413–426.

Monnahan, C., Branch, T.A. and A.E. Punt. 2015. Do ship strikes threaten the recovery of endangered eastern North Pacific blue whales? *Marine Mammal Science* 31: 279-297.

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

Mori, M. and D.S. Butterworth. 2006. A first step towards modelling the krill-predator dynamics of the Antarctic ecosystem. *CCAMLR Science* 13: 217-277.

Müller, A., Butterworth, D.S. and S.J. Johnston. 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. IWC Document SC/63/SH26. 45pp.

Nadeem, K., Moore, J.E., Zhang, Y. and H. Chipman. 2016. Integrating population dynamics models and distance sampling data: a spatial hierarchical state-space approach. *Ecology* 97: 1735-1745.

Punt, A.E 1999. Report of the Scientific Committee. Annex R. A full description of the standard BALEEN II model and some variants thereof. *Journal of Cetacean Research and Management* (*Supplement*) 1: 267-276.

Punt, A.E. 2006. Assessing the Bering-Chukchi-Beaufort Seas stock of bowhead whales using abundance data together with data on length or age. *Journal of Cetacean Research and Management* 8: 127-137.

Punt, A.E. 2015a. Initial evaluation of two options for addressing infrequent surveys of the Bering-Beaufort Seas bowhead whales. IWC Document SC/D15/AWMP1. 21pp.

Punt, A.E. 2015b. An investigation of the relationship between $MSYR_{MAT}$ and $MSYR_{1+}$ based on a non-individual-base model. *Journal of Cetacean Research and Management* (*Supplement*) 16: 110-111.

Punt, A.E. 2016. A revised age-structured model for exploring the conceptual models developed for gray whales in the North Pacific. IWC. Document SC/A16/GW2. 24pp.

Punt, A.E. and J.M. Breiwick. 2002. A framework for evaluating Strike Limit Algorithms for population reduced to small numbers. *Journal of Cetacean Research and Management* 4: 165-177.

Punt, A.E. and D.S. Butterworth. 1997. Assessments of the Bering-Chukchi-Beaufort Seas stock of bowhead whales (*Balaena mysticetus*) using maximum likelihood and Bayesian methods. *Reports of the International Whaling Commission* 47: 603-618.

Punt, A.E. and G. Donovan. 2007. Developing management procedures that are robust to uncertainty: Lessons from the International Whaling Commission. *ICES Journal of Marine Science* 64: 603-612.

Punt, A.E. and T. Polacheck. 2008. Further analyses related to application of statistical catch-at-age analysis to Southern Hemisphere minke whales. IWC Document SC/60/IA2. 46pp.

Punt, A.E. and P.R. Wade. 2012. Population status of the eastern North Pacific stock of gray whales in 2009. *Journal of Cetacean Research and Management* 12: 15-28.

Punt, A.E., Butterworth, D.S., de Moor, C.L., De Oliveira, J.A.A. and M. Haddon. 2016. Management Strategy Evaluation: Best Practices. *Fish and Fisheries* 17: 303-334.

Punt, A.E., Friday, N.A. and T.D. Smith. 2006. Reconciling data on the trends and abundance of North Atlantic humpback whales within a population modelling framework. *Journal of Cetacean Research and Management* 8: 145-159.

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

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

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

Reilly, S.B., Rice, D.W. and A.A. Wolman. 1983. Population assessment of the gray whale, *Eschrichtius robustus*, from California shore censuses, 1967-80. *Fishery Bulletin* 81: 267-268

Ross-Gillespie, A., Butterworth, D.S. and K. Findlay. 2014. Assessment results for humpback breedings stocks D, E1 and oceania following recommendations from SC 65a. IWC Document SC/65b/SH04Rev. 28pp.

Ross-Gillespie, A., Butterworth, D.S. and K. Findlay. 2015. Final results for the final 'base case' three-stock DSB, BSE1 and BSO model, with sensitivity runs. *Journal of Cetacean Research and Management* (*Supplement)* 16: 215-221.

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

Shirakihara, K. and S. Tanaka. 1984. Simulation studies on length-specific population assessment techniques for western North Pacific sperm whales. *Reports of the International Whaling Commission* 34: 259-264..

Shirakihara, K., Tanaka, S. and T. Nakano. 1983. The revised age-specific model for population assessment of the Western North Pacific sperm whales. *Reports of the International Whaling Commission* 33: 757-759.

Shirakihara, K., Tanaka, S. and T. Nakano. 1985. Further simulation studies on length-specific population assessment techniques for western North Pacific sperm whales. *Reports of the International Whaling Commission* 35: 199-203.

Siler, W. 1979. A competeing-risk model for animal mortality. *Ecology* 64: 750–757.

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 Padua. vi+103pp,

Slooten, E. 2015. Effectiveness of partial protection for Maui's dolphin. IWC Document SC/66a/SM12. 8pp.

Smith, T.D. and R.R. Reeves. 2003. Estimating historic humpback removals from the North Atlantic: an update. *Journal of Cetacean Research and Management* (*Supplement*) 5: 301–311.

Stolen, M.K. and J. Barlow. 2003. A model life table for bottlenose dolphins (*Tursiops truncates*) from the Indian River lagoon system Florida, U.S.A. *Marine Mammal Science* 19: 630-649.

Tilman, M.F. 1977. Estimates of population size for the North Pacific sei whale. *Reports of the International Whaling Commission* (Special Issue 1): 98-106.

Towell, R. 2007. Population Dynamics of Northern Fur Seals on the Pribilof Islands, Alaska. MS Thesis, University of Washington.

Wade, P.R. 2002. A Bayesian stock assessment of the eastern Pacific gray whale using abundance and harvest data from 1967-1996. *Journal of Cetacean Research and Management* 4: 85-98.

Wade, P.R., Reilly, S.B. and T. Gerrodette. 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 S.B. Reilly. 2007. Depletion of spotted and spinner dolphins in the eastern tropical Pacific: modeling hypotheses for their lack of recovery. *Marine Ecology Progress Series* 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 M.P. Heidi-Jørgensen. 2015a. Population model for West Greenland beluga. NAMMCO Document. NAMMCO/SC/22/JCNB/SWG/2015-JWG/09. 12pp.

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

Zerbini, A.N., Clapham, P.J. and P. R. Wade. 2010. Assessing plausible rates of population growth in humpback whales from life-history data. *Marine Biology* 157: 1225–1236.

Zerbini, A.N., Ward, E.J., Kinas, P.G., Engel, M.H. and A. Andriolo. 2011. A Bayesian assessment of the conservation status of humpback whales (*Megaptera novaenangliae*) in the western South Atlantic Ocean (Breeding Stock A). *Journal of Cetacean Research and Management* (Special Issue) 3: 131-144.

Zeh, J.E. and A.E. Punt. 2005. Updated 1978-2001 abundance estimates and their correlations for the Bering-Chukchi-Beaufort Seas stock of bowhead whales. *Journal of Cetacean Research and Management* 7: 169-175.

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)<br>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 & 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 (Western Australia) + | | | | Ross Gillespie *et al*. (2014, 2015) |
| BSE (Eastern Australia) + | Assessment | Production model | 3 stock model | Jackson *et al.* (2016) |
| BSF (Oceania) | | | | |
| BSG (West South America) | Assessment | Production model | 1 stock | Johnston *et al*. (2011) |
| **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);<br>MSE (d) | Production model (a)<br>Sex- and age-structured (b,d)<br>Sex-, age- and stage-structured (c) | 1 stock | Wade (2002) (a);<br>Punt and Wade (2012) (b)<br>Brandon and Punt (2013)(c)<br>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) | 1 stock | Wade *et al*. (2002) (a); |
| | | Production model (a,b,d,e) | | Wade *et al*. (2007) (b); |
| | | Age- sex-, and stage-structured (c) | | 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) | 1 stock | Wade *et al*. (2002) (a); |
| | | Production model (a,b,c,d) | | Wade et al. (2007) (b); |
| | | Age-structured model (a) | | Lennert-Cody *et al*. (2012) (c); Anon (2009) (d) |
| *Whitebelly* | | Production model | 1 stock | Anon (2009) |
| **Common bottlenosed dolphin** | | | | |
| *Northern Adriatic* | Assessment | Stochastic production model | 1 stock | Simeoni (2014) |
| **Beluga 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) |
| **Harbor Porpoise** | | | | |
| *West Greenland* | Assessment | Sex- and age-structured | 1 stock | Witting (2013) |

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 whales** | | | | | |
| *Indian & 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 heterogeneity / 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)<br>No (b,c,d,e) | No (a,b,d,e)<br>Yes (c) | No |
| *Southern/western* | | | | | |
| **Spinner dolphin** | Yes | Yes | No | No | No |
| *Eastern* | Yes | Yes | Yes (a)<br>No (b,c,d) | No | No |
| *Whitebelly* | Yes | Yes | No | No | No |
| **Common bottlenosed 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 |
| **Harbor Porpoise** | Yes | Yes | Yes | No | No |
| *West Greenland* | Yes | Yes | No | Yes | No |

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)<br>Births (Pella-Tomlinson form) (a,b,c); estimated MSYR (a,b) Stochastic recruitment (c)[&] | Age-specific[$] (estimated) (a,b,c); episodic mortality events (c)[&] | Pre-specified (a,c,d);<br>Estimated (b) | Brandon and Wade (2006) examined several alternative models. |
| West Greenland | Births (Pella-Tomlinson form); pre-specified MSYR (stochastic recruitment[&]) | Pre-specified / episodic mortality events[&] | Pre-specified | |
| **Southern right whale** | | | | |
| *Southeast Atlantic* | No | Age-specific[$] (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 whales** | | | | |
| *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 a 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);<br>• estimated MSYR (a,b);<br>• estimated MSYR with stochastic recruitment (c)<br>• 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 | Pre-specified | Diffusion (estimated); Effect of mass mortality event (estimated) |

| | | | | |
|---|---|---|---|---|
| | | events[&] | | |
| 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; * But 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 bottlenosed 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 |
| **Beluga 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) (parameterized 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 | |
| **Harbor Porpoise** | | | | |
| *West Greenland* | Births (Pella-Tomlinson form) & exponential growth (productivity estimated) | Estimated (age-specific) | Estimated | Assessment not used for management |

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)<br>Yes (b, c, d) | K, MSY, MSYR (a,b); K, MSYL (c,d) | RY, $N^{tot}$ (s)<br>$N^{tot}$, need satisfaction, recovery rate, catch variation (b,c,d6) | 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 whales** | | | | |
| *Indian & 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 whales** | | | | |
| *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 |

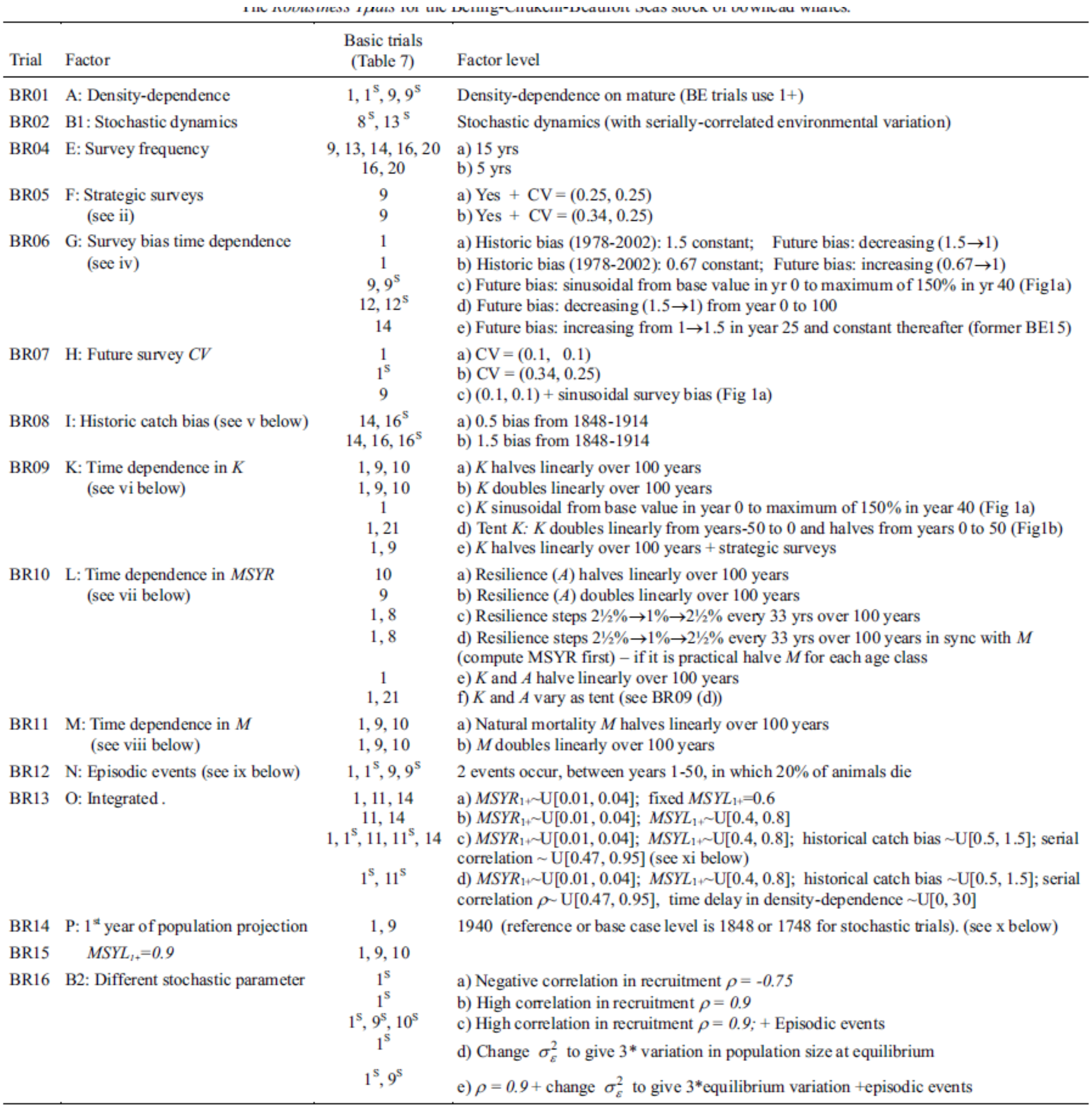
| | | | | |
|---|---|---|---|---|
| **Gray whale** | | | | |
| *North Pacific* | | | | |
| Western gray whale | Yes | | $N^{tot}$, $N^{mat}$, Immigrants | Bayesian |
| Eastern gray whale | Yes (a,c,d)) | MSY,RY (a); K (b,c,d) | $N^{tot}$ (a,b); Ntot, need satisfaction (c); $N^{tot}/K^{tot}$; recovery 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 bottlenosed dolphin** | | | | |
| *Northern Adriatic* | No | K | $N^{ot}$ | Bayesian |
| **Beluga whales** | | | | |
| *Eastern Hudson Bay* | Yes (catches) | None | $N^{tot}$ | Bayesian |
| *Cook Inlet* | Yes | Extinction | $N^{tot}$ | Bayesian |
| *West Greenland* | | | | |
| **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 |
| **Harbor Porpoise** | | | | |
| *West Greenland* | Yes (catches) | RY, K | $N^{tot}$ | Bayesian |

Table 9. The Robustness trials (sensitivity test) for the MSE for the Bering-Chukchi-Beaufort Seas bowhead whales (from IWC, 2003)



The *Robustness Trials* for the Bering-Chukchi-Beaufort Seas stock of bowhead whales.

| Trial | Factor | Basic trials (Table 7) | 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<br>16, 20 | a) 15 yrs<br>b) 5 yrs |
| BR05 | F: Strategic surveys (see ii) | 9<br>9 | a) Yes + CV = (0.25, 0.25)<br>b) Yes + CV = (0.34, 0.25) |
| BR06 | G: Survey bias time dependence (see iv) | 1<br>1<br>$9, 9^S$<br>$12, 12^S$<br>14 | a) Historic bias (1978-2002): 1.5 constant; Future bias: decreasing (1.5→1)<br>b) Historic bias (1978-2002): 0.67 constant; Future bias: increasing (0.67→1)<br>c) Future bias: sinusoidal from base value in yr 0 to maximum of 150% in yr 40 (Fig1a)<br>d) Future bias: decreasing (1.5→1) from year 0 to 100<br>e) Future bias: increasing from 1→1.5 in year 25 and constant thereafter (former BE15) |
| BR07 | H: Future survey *CV* | 1<br>$1^S$<br>9 | a) CV = (0.1, 0.1)<br>b) CV = (0.34, 0.25)<br>c) (0.1, 0.1) + sinusoidal survey bias (Fig 1a) |
| BR08 | I: Historic catch bias (see v below) | $14, 16^S$<br>$14, 16, 16^S$ | a) 0.5 bias from 1848-1914<br>b) 1.5 bias from 1848-1914 |
| BR09 | K: Time dependence in *K* (see vi below) | 1, 9, 10<br>1, 9, 10<br>1<br>1, 21<br>1, 9 | a) *K* halves linearly over 100 years<br>b) *K* doubles linearly over 100 years<br>c) *K* sinusoidal from base value in year 0 to maximum of 150% in year 40 (Fig 1a)<br>d) Tent *K: K* doubles linearly from years-50 to 0 and halves from years 0 to 50 (Fig1b)<br>e) *K* halves linearly over 100 years + strategic surveys |
| BR10 | L: Time dependence in *MSYR* (see vii below) | 10<br>9<br>1, 8<br>1, 8<br>1<br>1, 21 | a) Resilience (*A*) halves linearly over 100 years<br>b) Resilience (*A*) doubles linearly over 100 years<br>c) Resilience steps 2½%→1%→2½% every 33 yrs over 100 years<br>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<br>e) *K* and *A* halve linearly over 100 years<br>f) *K* and *A* vary as tent (see BR09 (d)) |
| BR11 | M: Time dependence in *M* (see viii below) | 1, 9, 10<br>1, 9, 10 | a) Natural mortality *M* halves linearly over 100 years<br>b) *M* doubles linearly over 100 years |
| BR12 | N: Episodic events (see ix below) | $1, 1^S, 9, 9^S$ | 2 events occur, between years 1-50, in which 20% of animals die |
| BR13 | O: Integrated . | 1, 11, 14<br>11, 14<br>$1, 1^S, 11, 11^S, 14$<br>$1^S, 11^S$ | a) $MSYR_{1+}$~U[0.01, 0.04]; fixed $MSYL_{1+}$=0.6<br>b) $MSYR_{1+}$~U[0.01, 0.04]; $MSYL_{1+}$~U[0.4, 0.8]<br>c) $MSYR_{1+}$~U[0.01, 0.04]; $MSYL_{1+}$~U[0.4, 0.8]; historical catch bias ~U[0.5, 1.5]; serial correlation ~ U[0.47, 0.95] (see xi below)<br>d) $MSYR_{1+}$~U[0.01, 0.04]; $MSYL_{1+}$~U[0.4, 0.8]; historical catch bias ~U[0.5, 1.5]; serial correlation $\rho$~ U[0.47, 0.95], time delay in density-dependence ~U[0, 30] |
| BR14 | P: $1^{st}$ year of population projection | 1, 9 | 1940 (reference or base case level is 1848 or 1748 for stochastic trials). (see x below) |
| BR15 | $MSYL_{1+}=0.9$ | 1, 9, 10 | |
| BR16 | B2: Different stochastic parameter | $1^S$<br>$1^S$<br>$1^S, 9^S, 10^S$<br>$1^S$<br>$1^S, 9^S$ | a) Negative correlation in recruitment $\rho = -0.75$<br>b) High correlation in recruitment $\rho = 0.9$<br>c) High correlation in recruitment $\rho = 0.9$; + Episodic events<br>d) Change $\sigma_\varepsilon^2$ to give 3* variation in population size at equilibrium<br>e) $\rho = 0.9$ + change $\sigma_\varepsilon^2$ to give 3*equilibrium variation +episodic events |

`Table 10. Tentative best practice guidelines for cetacean stock assessments

| Issue | Guidelines |
|---|---|
| **Model structure assumptions** | |
| Spatial and stock structure | Required if genetic (or non-genetic) 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 | Generally unnecessary, but can be used to impose assumptions regarding calving intervals. |
| 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, but a later year 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-specific (minimally calf, non-calf; but alternative forms such as the Siler form should be considered). |
| Selectivity | Generally 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 summarized data | Ideally, 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 analyzing the raw data and those underlying the population to 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 actual application |

## Appendix A: Thoughts related to Dolphin stocks in the eastern tropical Pacific Ocean

The ideal assessment model for dolphin stocks in the eastern tropical Pacific Ocean depends (as is the case for all analyses) on the aims of modelling. If, for example, the aim is understand recent trends in abundance, a log-linear model (implemented as state-space model) would seem the best way to proceed. However, an assessment would need to be based on a population model (ideally age- and sex-structured) if status relative to reference points such as the carrying capacity or MSYL was required. Herewith is my view of how past assessments for these dolphins stocks (with a focus on those for eastern spotted dolphin and eastern spinner dolphin) follow each guideline in Table 10. In general, the approach of Hoyle and Maunder (2004) comes closest to following each guideline:

- Stock structure: the documents describing the assessments do not very clearly specify the basis for the stock structure hypothesis on which the assessment is based.
- The models include age-, sex- and stage-structure.
- The stock was assumed to be at carrying capacity prior to removals, but the removals are fairly uncertain, particularly during the early years, and this uncertainty is not reflected in the analysis, either directly or through sensitivity analyses.
- Only environmental stochasticity was considered, which seems appropriate given the size of the stock.
- Natural mortality was assumed to be independent of sex and age, but with no plus group. The model implicitly allowed for calf survival to differ from that of non-calves given density-dependent calf survival was included in the model. Dropping the assumption regarding the lack of a plus-group changed several of the parameter estimates.
- Selectivity was estimated and assumed to be logistic.
- Deviations about the “stock-recruitment relationship” were considered, but unlike Punt *et al*. (2014) no maximum was imposed on the calving rate / calf survival rate. Moreover, the extent of variation about the “stock-recruitment relationship” was assumed rather than being estimated even though this could have been possible given the analysis was Bayesian. The deviations appear strongly auto-correlated, suggesting either model mis-specification or that calf survival is related to some auto-correlated environmental (or fishery) effect.
- Most of the prior distributions were uniform, with exception of those for the recruitment deviations, and the parameter that determines the shape of the production function ($z$), which was taken to be normal. However, $z$ was pre-specified when sampling from the posterior distribution (as was the age-at-maturity).
- The assessment was Bayesian, but only a limited set of sensitivity analyses were undertaken, and no simulation evaluation of the estimation approach appears to have been performed.