

Defining the appropriate ‘Unit-To-Conserve’ under the International Whaling Commission’s Revised Management Procedure

K.K. MARTIEN¹, D.P. GREGOVICH¹ AND A.E. PUNT²

Contact e-mail: Karen.Martien@noaa.gov

ABSTRACT

Identifying the appropriate ‘Unit to Conserve’ (UTC) is critical to the success of any management scheme. While the need to define the UTC appropriate to the IWC’s Catch Limit Algorithm (CLA) has long been recognised by its Scientific Committee, little progress has been made on this issue. The CLA was rigorously tested prior to its adoption. However, most of those original performance trials focused on single-population scenarios or two-population scenarios with no ongoing dispersal. None of the trials considered the performance of the CLA across a range of dispersal rates. In this study, the performance of the CLA under a variety of population structure scenarios is examined. This is the first study to investigate the levels of connectivity (i.e. dispersal rate) for which populations require separate management to meet the conservation goals of the CLA. All the trials consisted of two populations that were managed as a single stock for 100 years. Both historical and modern hunts were spatially-biased so that population 1 was the primary target of hunting. Parameters that varied among trials were the relative carrying capacities (K) of the populations, the dispersal rate between them, maximum sustainable yield rate ($MSYR_{1+}$), and the precision in simulated abundance estimates. All of these parameters had strong effects on the conservation performance of the CLA. Trials with a low $MSYR_{1+}$ (1%) generally ended with the abundance of population 1 below $0.54K$, regardless of the dispersal rate or relative carrying capacities of the two populations. The same was true of trials in which the carrying capacity of population 1 represented only 10% of the total landscape carrying capacity and the CV of the abundance estimates was low, even when dispersal between populations was high ($5 \times 10^{-3} \text{yr}^{-1}$) and $MSYR_{1+}$ was 4%. The results suggest that the appropriate UTCs under the RMP are likely to exchange dispersers at high enough rates that they will be difficult to delineate using existing methods. These results also highlight the value of spatially-diffuse hunting patterns that avoid potential overhunting of unrecognised stocks.

KEYWORDS: CONSERVATION; UNIT TO CONSERVE; MANAGEMENT PROCEDURE; SUSTAINABILITY; STOCK IDENTITY

INTRODUCTION

The success of most management schemes hinges on accurate identification of the management units the scheme is intended to conserve. The dispersal rates at which ‘Units to Conserve’ (UTCs) should be defined will depend on the management objectives those units are intended to meet (Palsbøll *et al.*, 2007; Taylor, 2005; Taylor *et al.*, 2010; Waples and Gaggiotti, 2006). For instance, in the United States, the two major pieces of legislation aimed at conserving marine mammals are the Endangered Species Act (ESA), and the Marine Mammal Protection Act (MMPA). These Acts have different goals and hence require different levels of connectivity between their respective UTCs; the ESA aims to prevent the extinction of species, while the MMPA aims to maintain populations as ‘functioning elements of their ecosystems’. To achieve the goal of the ESA requires identification and management of units that contribute to the evolutionary potential of the species (Waples, 1991). Units that experience gene flow at the rate of one disperser or fewer per generation have been suggested as meeting this criterion because such a low level of gene flow could allow for the development of local adaptations (Gardenfors *et al.*, 2001; Taylor, 2005; Taylor *et al.*, 2010). Units with this level of connectivity are typically referred to as ‘Evolutionarily Significant Units,’ or ESUs (Moritz, 1994; Waples, 1991). The more ambitious conservation goal of the MMPA, on the other hand, requires separate management if

two units are demographically independent (Angliss and Wade, 1997). Specifically, simulations have shown that separate management is necessary even for populations exchanging dispersers at rates in excess of 1% per year if the goals of the MMPA are to be met (Taylor, 1997). Units at this level are referred to as ‘Demographically Independent Populations,’ or DIPs (Taylor, 1997; Waples and Gaggiotti, 2006).

The IWC agreed to use the Revised Management Procedure (RMP) to manage commercial whaling of baleen whales on their feeding grounds, should commercial whaling be allowed to resume (IWC, 1994). Catches are determined by the Catch Limit Algorithm (CLA), which was subjected to extensive performance testing to ensure robustness considering the three competing management objectives of maximising catch, minimising variability in catch and adequately conserving populations (IWC, 1991; 1992b; 2007). During the development of the CLA, testing was done to examine performance when two populations were erroneously managed as one stock (‘stock’ is the name given to management units in the RMP framework) and when one stock is managed as two (Smith *et al.*, In press). The two-population trials included scenarios in which dispersal (i.e. permanent movement of individuals from one breeding population to another) occurred between adjacent populations at a rate of 0.5% per year as well as cases with no dispersal. These trials confirmed that the conservation

¹ Southwest Fisheries Science Center, NMFS, NOAA, 8901 La Jolla Shores Drive, La Jolla, CA 92037, USA

² School of Aquatic and Fishery Sciences, University of Washington, Box 355020, Seattle, WA 98195-5020, USA

performance of the CLA is poor when two populations are erroneously combined into a single stock and hunting is concentrated on only one of the populations (IWC, 1991), but that stock definition errors are partially mitigated by dispersal between populations (IWC, 1992a). However, the trials did not investigate a range of dispersal rates to determine the critical level of dispersal above which the conservation goal of the RMP can be met even if two populations are combined into a single stock, i.e. they did not provide insight into the question of what is the appropriate UTC under the RMP. The RMP includes options (variants) such as catch cascading and catch capping to account for stock structure uncertainty, and evaluations of the consequences of stock definition errors have been conducted as part of selecting among these variants for individual stocks based on case-specific *Implementation Simulation Trials* (e.g. IWC, 2004; 2007; Punt and Donovan, 2007). However, the case-specific nature of *Implementation Simulation Trials* limits their utility for drawing general conclusions regarding the circumstances under which management as separate stocks is necessary to adequately protect populations.

In this study the performance of the CLA under a variety of population structure scenarios is examined. Two populations are erroneously managed as a single stock in all of our simulations. Performance is measured as the probability that the abundance of both populations is greater than 54% of carrying capacity (i.e. $0.54K$) by the end of each simulation. $0.54K$ was the Protection Level in the New Management Procedure, and the probability of taking whales when a stock was below $0.54K$ was taken into consideration when selecting among CLA variants. The performance of the CLA is evaluated as a function of the relative sizes of the two populations and the dispersal rate between them. Our results provide insight into the critical level of dispersal above which two populations can be managed as a single stock without compromising the conservation goal of the RMP. The results will provide guidance to researchers attempting to define stocks for individual species being managed under the RMP, as well as to those seeking to develop new approaches to defining stocks for use under the RMP.

METHODS

The TOSSM package (Martien *et al.*, 2009) was used to evaluate the performance of the CLA in the face of unrecognised population structure. A series of simulations were conducted in which two populations were combined into a single management unit (MU). A single catch limit was calculated by the CLA for the MU and was therefore based on the combined abundance of both populations. The entire catch limit was removed each year and, if possible, was taken entirely from 'population 1'. If there were insufficient individuals in population 1 to meet the catch limit, the balance of the catch limit was taken from population 2. Hunted animals were chosen at random, without respect to age or sex.

All simulations were initialised with both populations at K . The combined carrying capacity of the two populations (K_{TOT}) was set to 7,500 in all simulations. In the first year of

each simulation, population 1 was depleted to $0.3K$ and population 2 to $0.99K^3$. The two populations then underwent 100 years of simulated management, with both populations combined into a single MU. Abundance in the MU was estimated and the catch limit re-calculated every five years during this management period. We performed 100 replicates of each simulation.

Preliminary analyses showed that our results were strongly dependent on the coefficient of variation (CV) of the abundance estimate for the MU. In the TOSSM package, CV is modeled according to the formula

$$CV = 0.1\sqrt{x/n}$$

where N is the estimated abundance of the MU and x is the abundance at which $CV = 0.1^4$.

Taylor *et al.* (2007) found that CVs for stocks of large baleen whales in US territorial waters ranged from 0.08 (western North Pacific humpback whales) to 0.73 (CA/OR/WA minke whales), with an average of 0.29. Values of x were chosen such that when both populations in the MU were at carrying capacity, the CV was 0.30, 0.17, or 0.04 (Fig. 1). The highest of these values represents a typical level of precision, while the middle value represents a realistic high precision case⁵. The lowest CV examined is unlikely to apply or to be achieved for large whale species except in cases of mark-recapture estimates with high effort, but allows full characterisation of the behavior of the CLA. Higher values of CV were not examined. Doing so would have resulted in adequate protection under virtually all population structure scenarios because the CLA reduces catches when abundance estimates are imprecise, and would therefore not have contributed to the objective of identifying the population structure scenarios under which the CLA failed to adequately protect populations if population structure went undetected.

TOSSM simulations must be initialised with a dataset generated by the R package 'rmetasim' (Strand, 2002). An array of datasets with various parameterisations were generated using 'rmetasim' for this study. The three parameters varied in these datasets were: (1) the maximum sustainable yield rate ($MSYR_{+}$); (2) the carrying capacities of the two populations (always summing to 7,500); and (3) the annual dispersal rate between the two populations. A full description of all population parameters used when generating these datasets is included in Appendix 1.

Two main sets of trials were performed. The first set of trials was performed to determine the effect of dispersal rate on conservation performance, which we defined as recovering the most vulnerable population to $0.54K^6$. In these trials, the carrying capacity of each population was set at $K = 3,750$ (half of the total for the landscape). These trials

³These values were chosen to be consistent with those used during development and testing of the CLA (IWC, 1991)

⁴This is a simplification of the way CVs are generated by the control program used to test the CLA (IWC, 1991), but captures the essence of that data generation process.

⁵The guidelines to the RMP impose a minimum CV of 0.15 because performance of the CLA can be poor when the CV is underestimated (IWC, 1992b).

⁶Recovery to $0.54K$, which is approximately $MSYL$, was a measure of conservation performance used during the development of both the CLA and the algorithm for calculating catch limits under the MMPA.

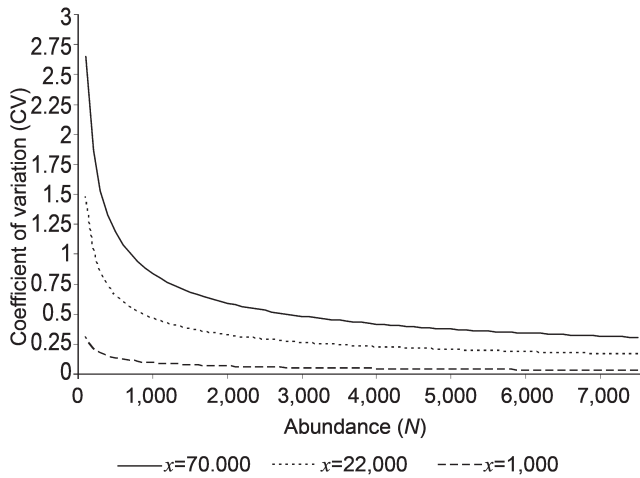


Fig. 1. Relationship between abundance and CV for the three values used for the CV tuning parameter (x). The values chosen for x result in CVs at $N = 7,500$ of 0.3 ($x = 70,000$), 0.17 ($x = 22,000$) and 0.04 ($x = 1,000$).

were conducted using the highest CV scenario ($x = 70,000$; Fig. 1). Four dispersal rates ($5 \times 10^{-6}\text{yr}^{-1}$, $5 \times 10^{-5}\text{yr}^{-1}$, $5 \times 10^{-4}\text{yr}^{-1}$, and $5 \times 10^{-3}\text{yr}^{-1}$) were examined. These rates span the range of dispersal rates typically of interest when defining units to conserve (Palsbøll *et al.*, 2007; Taylor *et al.*, 2010). Simulations using these four dispersal rates were performed for $\text{MSYR}_{1+} = 4\%$ and $\text{MSYR}_{1+} = 1\%$.

The second set of trials explored the effect of the CV of the abundance estimates on conservation performance across five levels of relative population carrying capacity. In these trials, the parameters held constant were dispersal rate ($5 \times 10^{-3}\text{yr}^{-1}$) and MSYR_{1+} (4%). For each of the three CV curves shown in Fig. 1, the relative carrying capacities of the two populations were varied so that population 1 comprised 10, 20, 30, 40, or 50% of the total landscape carrying capacity.

For each trial, the abundance trajectory for population 1 and the total catch during RMP hunting was summarised. Only the abundance of population 1 is presented because a spatial bias in hunting resulted in this population being more heavily impacted by hunting than population 2. This spatial bias simulated a situation in which whalers attempt to minimise effort by concentrating their effort close to a home port. Population 1 is assumed to be a coastal population and therefore the first encountered when whalers leave port. Population 2 is adjacent to population 1, but further offshore. The spatial bias in hunting during the historic hunting phase of the simulations is reflected in the initial depletion of population 1 (it is depleted to 0.3K prior to the first application of the CLA). The extent of spatial bias in removals during the management period is controlled in the TOSSM package by the ‘harvest.interval’ argument. Hunting always occurs initially in the harvest interval closest to the coast and then proceeds steadily offshore upon depletion of the animals in each interval (Martien *et al.*, 2009). For all of our simulations the landscape was divided into ten equally-sized harvest intervals.

Two single-stock simulations were ran and the results were compared with those from previous analyses to confirm that the CLA as implemented in the TOSSM package was working correctly. These trials each consisted of a single population with a carrying capacity of 7,500 that was initially

depleted to 0.3K. MSYR_{1+} was set at 1% in one trial and 4% in the other. Thus, these trials correspond fairly closely to the R1 and R4 base-case trials used during the development of the CLA (IWC, 1991), except that MSYR refers to the 1+ rather than mature component of the population, MSYL does not occur at 0.6K for our trials but instead occurs at 0.518K and 0.547K for the $\text{MSYR}_{1+} = 1\%$ and 4% trials, respectively, and the abundance estimates are generated slightly differently⁷. These single-stock trials were run for 500 years, with median abundance (across 100 replicates) recorded at years 100 and 500.

RESULTS

Population 1 always recovered to greater than 0.54K (Fig. 2) after 100 years of CLA management when the initial carrying capacities of the two populations were equal to half of the total landscape K , and MSYR_{1+} was set at 4%. However, this was not the case when MSYR_{1+} was set at 1%. Population 1 did not generally recover in these trials. The exception to this was when the dispersal rate, d , equaled $5 \times 10^{-3}\text{yr}^{-1}$. This dispersal rate resulted in better conservation performance (Fig. 2), with population 1 recovering to above 0.54K in 69% of replicates. The abundance trajectories for the three lowest dispersal rates ($d = 5 \times 10^{-6}$ to 5×10^{-4}) were fairly consistent (Fig. 2). The depletion of population 1 in the $\text{MSYR}_{1+} = 1\%$ trials resulted in lower total catches than in the $\text{MSYR}_{1+} = 4\%$ trials (Table 1).

Table 1
Median cumulative catch as a function of MSYR_{1+} and dispersal rate.

| MSYR_{1+} | Dispersal rate | | | |
|--------------------|--------------------|--------------------|--------------------|--------------------|
| | 5×10^{-6} | 5×10^{-5} | 5×10^{-4} | 5×10^{-3} |
| 4% | 2,855 | 2,775 | 2,850 | 2,900 |
| 1% | 1,590 | 1,660 | 1,700 | 1,825 |

The CVs for the abundance estimates strongly affected the population trajectories, as did the relative carrying capacities of the two populations (Fig. 3). Population 1 did not recover to 0.54K with high probability with the lowest CV regardless of the relative sizes of the populations, except when the carrying capacity for population 1 was half of the total when this probability was 0.73 (Fig. 3, upper panels). In contrast, Population 1 did recover in most trials with the two higher CVs, except when population 1 constituted only 10% of the total landscape K initially. The impact of higher CVs on population recovery rates was not unexpected; the CLA sets the catch limit as the 40.2th percentile of a posterior distribution for the catch limit; greater uncertainty in abundance estimates thus results in a wider posterior distribution for the catch limit and hence a lower catch limit overall. The results in Fig. 3 are based on the most optimistic assumptions regarding MSYR_{1+} (4%) and dispersal rate ($5 \times 10^{-3}\text{yr}^{-1}$).

⁷In principle, 0.54K is 10% less than the nominal MSYL of 0.6K, suggesting that the threshold for defining management success should be 10% lower than the 0.518K and 0.547K. However, the linkage between 0.54K and MSYL (if not the entire rationale for the linkage) is no longer used explicitly in testing of management strategies.

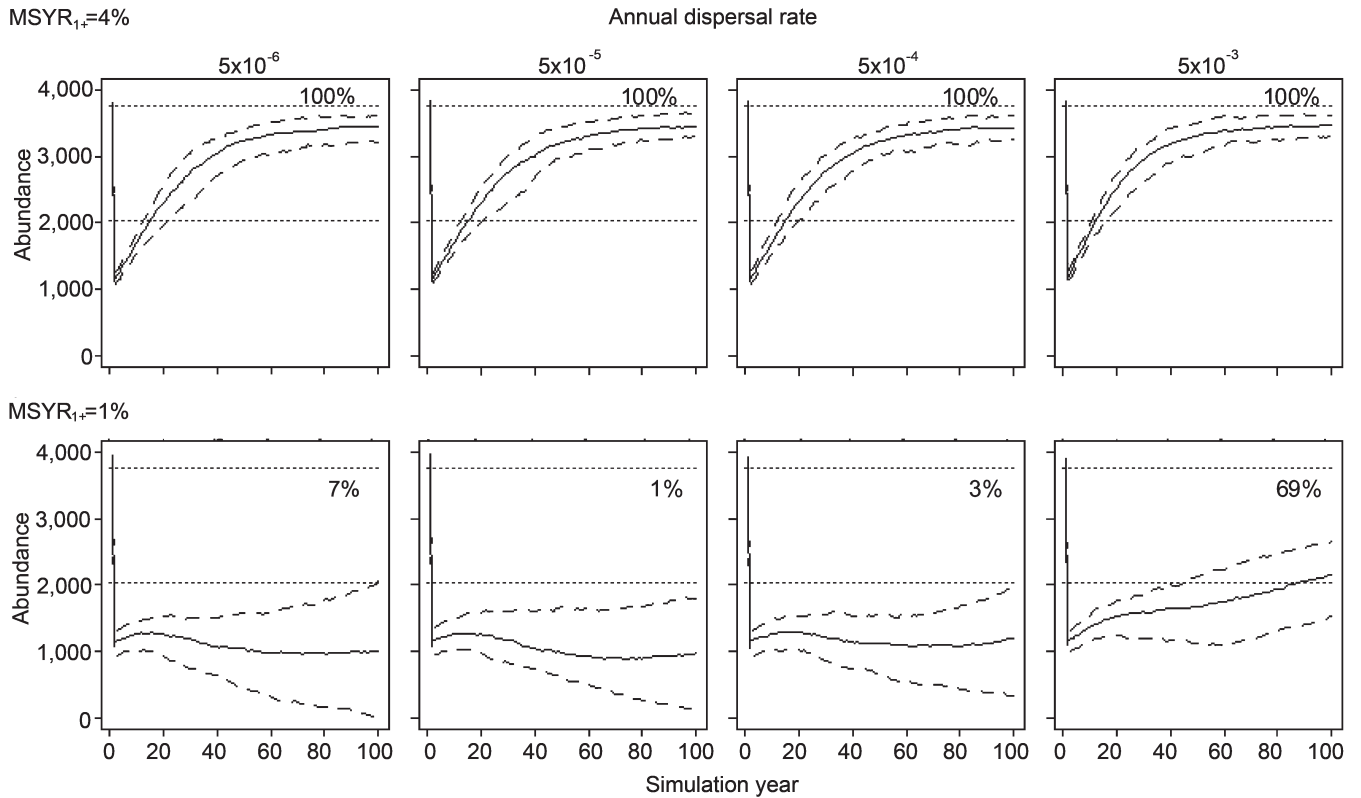


Fig. 2. Time-trajectories of total (1+) population size for population 1 (*PI*) as a function of $MSYR_{1+}$ (4% and 1%) and dispersal rate (ranging from 5×10^{-6} to $5 \times 10^{-3} \text{yr}^{-1}$). Median (solid line) and 5th and 95th percentiles (dashed lines) are shown. The carrying capacity for each population was 3,750 (upper dotted line), and the CV at carrying capacity was 0.3. The percentage of replicates whose final abundance was greater than 0.54K (lower dotted line) is shown on each panel.

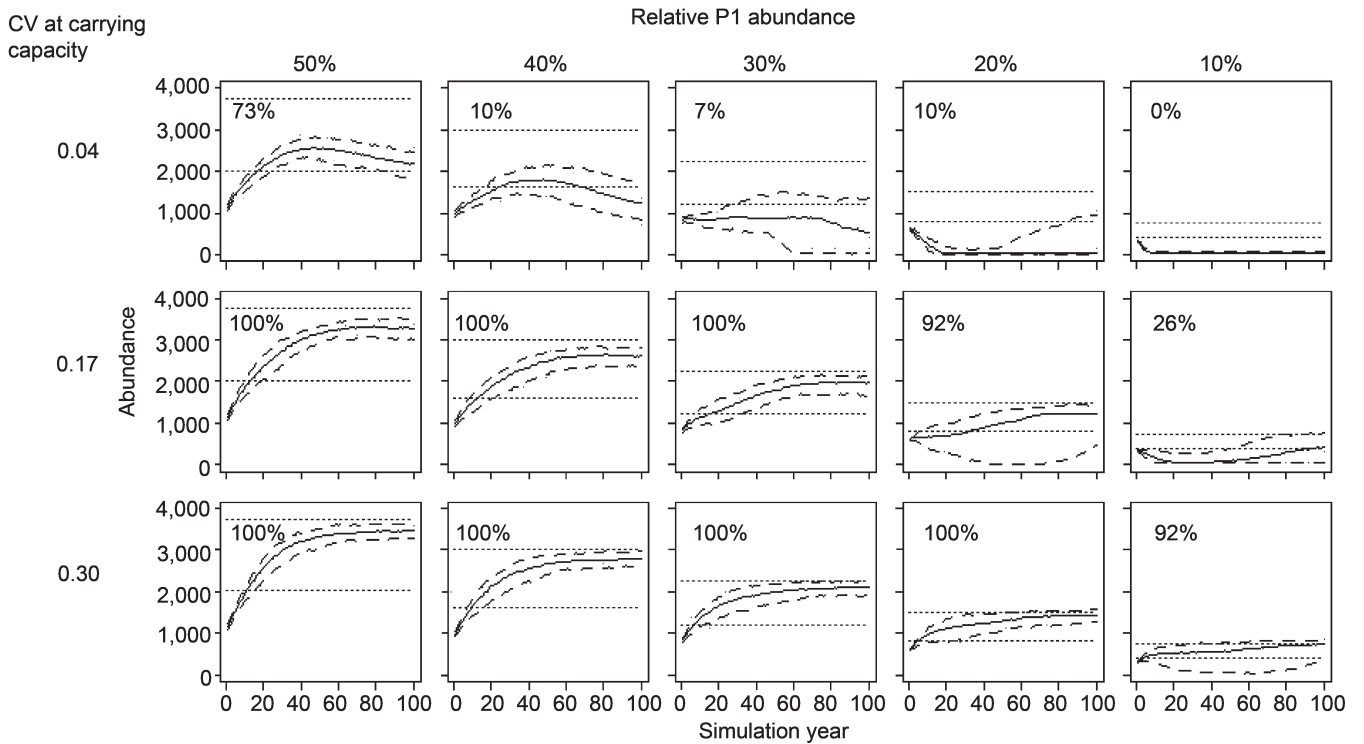


Fig. 3. Time-trajectories of total (1+) population size for population 1 (*PI*) as a function of the fraction which this population constitutes of the entire landscape carrying capacity and the CV at carrying capacity. Median (solid line) and 5th and 95th percentiles (dashed lines) of abundance are shown, along with K (upper dotted line) and $0.54K$ (lower dotted line). Dispersal rate and $MSYR_{1+}$ were set to $5 \times 10^{-3} \text{yr}^{-1}$ and 4%, respectively. The percentage of replicates whose final abundance was greater than 0.54K is shown on each panel. $MSYL$ for these simulations is 0.547K.

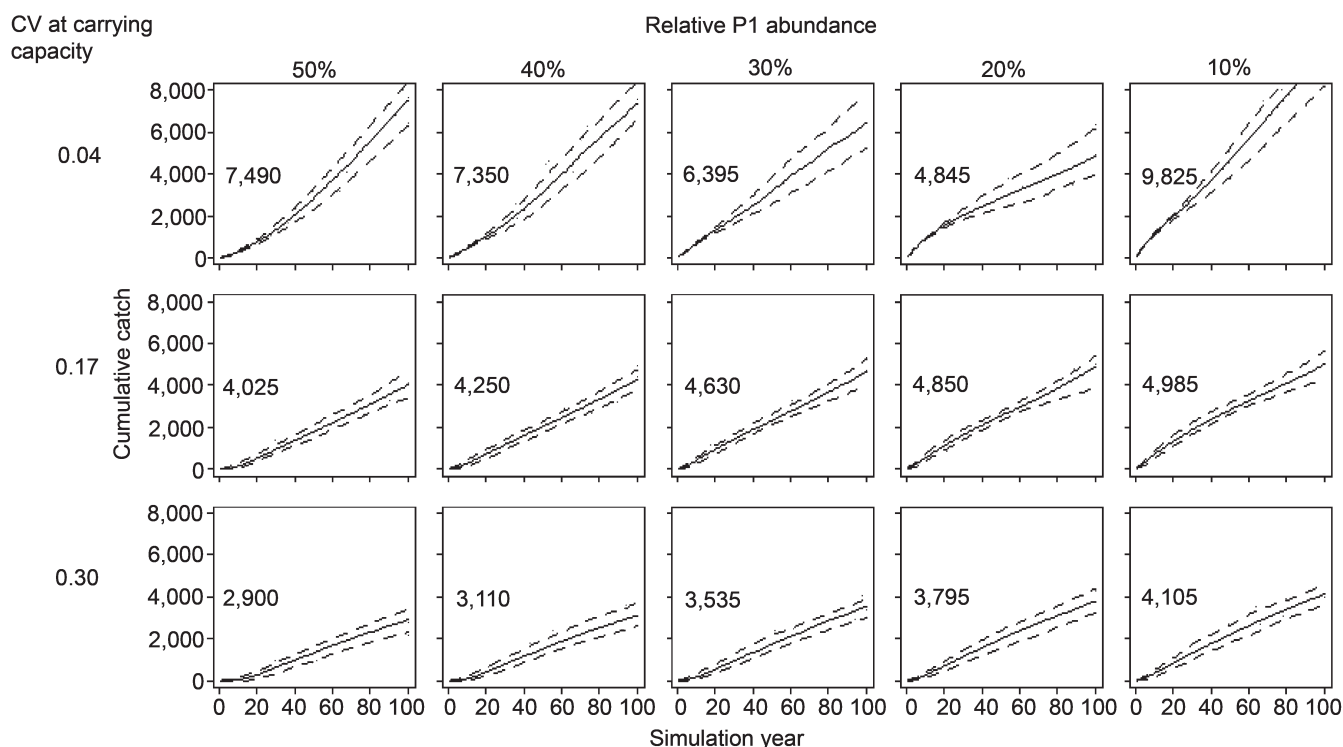


Fig. 4. Cumulative catch as a function of the fraction which population 1 constitutes of the entire landscape carrying capacity and the CV at carrying capacity (see Fig. 1). Median (solid line) and 5th and 95th percentiles (dashed lines) are shown. Dispersal rate and $MSYR_{1+}$ were held constant at $5 \times 10^{-3} \text{yr}^{-1}$ and 4%, respectively. Median total catch in year 100 is shown on each panel.

The total catch decreased as the CV of the abundance estimates increased (Fig. 4), as expected. Total catch increased as the relative initial abundance of population 1 decreased for the two higher CVs. The pattern was largely reversed in simulations using the lowest CV (cases in which population 1 fails to recover to $MSYL$), with total catch decreasing as the relative abundance of population 1 decreased from 0.5 to 0.2. However, the total catch more than doubled as the relative size of population 1 was further reduced from 20% to 10% of the landscape K .

The results of the single-stock trials were not inconsistent with those published by Butterworth and Punt (1994) (Table 2)⁸, indicating that the CLA is correctly implemented in the TOSSM package. Comparisons between our results and the graphical summaries of Aldrin *et al.* (2006) also provide support for the CLA working correctly within TOSSM.

DISCUSSION

Failure to manage populations separately resulted in poor conservation performance of the CLA for many of the parameter combinations examined, including some with dispersal as high as 0.5% per year, the highest value examined. Thus, it is necessary to separately manage populations between which annual dispersal rates are relatively high, at least from an evolutionary perspective, to protect and sustain populations of large whales. The fact that the conservation performance of the CLA is sensitive to errors in stock definition was demonstrated early in the development of the algorithm (IWC, 1991). However,

because performance across a range of dispersal rates was examined, this study provides new insights into the critical level of dispersal below which two populations warrant management as separate stocks. Our results indicate that the ‘unit-to-serve’ of relevance to the RMP is much more similar to the ‘Demographically Independent Populations’ (DIPs) of the MMPA than to the ‘Evolutionarily Significant Units’ (ESUs) of the ESA (Taylor, 2005; Taylor *et al.*, 2010). This is not surprising, given that the management objective of the RMP (i.e. maintaining sustainable fisheries) is focused on an ecological rather than an evolutionary scale (Waples and Gaggiotti, 2006).

Defining stocks that adequately protect populations managed under the RMP is likely to be challenging, especially if catches are not spread out spatially (e.g. using techniques such as catch cascading). In recent decades, genetic data have emerged as the most powerful tool available for defining units to conserve (Taylor *et al.*, 2010) and hence for identifying hypotheses for consideration when developing *Implementation Simulation Trials* (ISTs) for RMP testing. However, dispersal rates on the order of $5 \times 10^{-3} \text{yr}^{-1}$ result in very low levels of genetic differentiation. Most existing analytical methods are unable to detect such low levels of differentiation (Chen *et al.*, 2007; Latch *et al.*, 2006; Martien *et al.*, 2009; Martien and Taylor, 2003; Morin *et al.*, 2008; Taylor *et al.*, 2010; Waples and Gaggiotti, 2006). Accurately identifying stocks of large whales is made more difficult by the large abundance and long generation times of most species, both of which tend to reduce genetic differentiation for a given dispersal rate (Morin *et al.*, 2008). These factors also limit the power of non-genetic methods of investigating population structure, such as photo-identification and satellite telemetry. Our results therefore

⁸The results of Butterworth and Punt (1994) were based on $MSYR$ defined in terms of the mature rather than 1+ population component.

Table 2
Median of abundance at years 100 and 500 for single stock trials. The 5th and 95th percentiles are shown in parentheses.

| MSYR | TOSSM package | | Butterworth and Punt (1994) | |
|------|---------------------|--------------------|-----------------------------|-----------|
| | Year 100 | Year 500 | Year 100 | Year 500* |
| 1% | 0.702 (0.564-0.803) | 0.885 (0.627-1.00) | 0.624 (0.559-0.663) | 0.85 |
| 4% | 0.966 (0.932-1.00) | 0.971 (0.912-1.01) | 0.943 (0.843-0.981) | 0.85 |

*Butterworth and Punt (1994) only report median abundance for year 500 in graphical form. Thus, the values reported here are approximate.

emphasise the importance of giving careful consideration to even weak evidence suggesting population structure when developing plausible stock structure scenarios for *ISTs*.

The need to identify population structure accurately is particularly acute when $MSYR_{1+}$ is low. All but one of the $MSYR_{1+} = 1\%$ trials resulted in very poor (<0.10) conservation performance, and the probability of rebuilding to $0.54K$ was only 0.69 for the one case in which conservation performance was not very poor (Fig. 2). In contrast, the $MSYR_{1+} = 4\%$ trials examined proved relatively robust to errors in stock definition; population 1 failed to recover to $0.54K$ only when the *CV* was unrealistically low and in the trial in which population 1 comprised only 10% of the total landscape carrying capacity the *CV* at carrying capacity was 0.17, a low but realistic value.

The impact of errors in stock definition on the performance of the CLA is greater for stocks for which the *CV* of abundance estimates is low (Fig. 3). This interaction reflects the fact that an error in stock definition is, in essence, an error in estimating the geographic range, and therefore abundance, of the unit that is being affected by removals. The *CV* is the only input to the CLA that reflects uncertainty associated with the estimates of the abundance of the stock, although the CLA imposes a minimum *CV* for all abundance estimates, partially to reflect the fact that abundance estimates can contain sources of error not captured by sampling error alone. A high *CV* implies greater uncertainty regarding the number of animals available for hunt and can therefore partially compensate for abundance estimation errors due to mis-identification of stocks.

The *CV* also affects total catch, with catch decreasing with increasing *CV*. Fig. 4 shows a complicated interaction between *CV* and the relative carrying capacities of the two populations in determining total catch. For the two larger *CVs*, total catch increases as the size of population 1 decreases. This pattern reflects the fact that population 1 is depleted to $0.3K$ at the beginning of the simulations. Thus, simulations in which population 1 comprises a larger fraction of the overall carrying capacity have lower total abundances in the early years of the simulation, resulting in lower catch. For the lowest *CV*, however, the higher initial total abundance in simulations in which the relative size of population 1 is small is counteracted by the fact that population 1 continues to decline throughout the entire simulation due to overhunting (Fig. 3). This continued decline results in continual reductions in the catch limit, thus reducing total catch (Fig. 4). This pattern is dramatically reversed when the relative size of population 1 is reduced to

only 10%, which results in a more than doubling of total catch. In this scenario, population 1 is extirpated early in the simulation in all replicates. Thus, the landscape is managed 'correctly' for the majority of the simulation – a single (extant) population is managed as a single management unit.

The TOSSM package differs slightly from the model typically used in *ISTs* with respect to the way that *CV* is calculated. In the TOSSM package, the *CV* is inversely proportional to the total abundance. In contrast, *ISTs* (and the trials which were used to test the CLA) generate abundance estimates using *CVs* that include a constant term and a term that depends on the inverse of total abundance (e.g. IWC, 2004; IWC, 2007). The generation process includes both log-normal and Poisson components. Moreover, *ISTs* account for uncertainty caused by errors when estimating the sampling *CV* from a survey and often the impacts of temporal variation in migration of stocks. The TOSSM package is therefore less amenable to the type of case-specific, highly realistic simulations for which the models used in *ISTs* are designed. However, the differences between the TOSSM package and the *IST* models would not affect the conclusions of this study.

Our results show that extra caution is necessary to ensure that stocks have been correctly defined in cases where *CVs* of abundance estimates are low. Similarly, species with low *MSYRs* warrant special attention with respect to defining stocks, as do those cases in which adjacent populations differ markedly in abundance. It remains to be seen whether existing analytical methods are capable of identifying population structure at the level required for proper stock definition. Even if population structure can be accurately described, defining stocks that adequately protect populations can be very difficult in cases where populations overlap in areas where they are hunted, typically in the migration corridor or on feeding grounds. Managers should thus continue to employ methods to ensure spatially-diffuse hunting patterns (e.g. IWC, 2009) so as to reduce the risk of over-exploiting unrecognised populations, especially when power analysis indicates that available analytical methods would be unable to detect relevant levels of population structure.

ACKNOWLEDGMENTS

We would like to thank Mark Bravington, Barbara Taylor, and Eric Archer for their input into the design of this study. Many thanks also to Barbara Taylor, Tim Gerrodette, Eric Archer and two anonymous reviewers for their insightful comments on this paper.

REFERENCES

Aldrin, M., Storvik, B. and Schweder, T. 2006. Standardized catch per unit effort in minke whaling in Norwegian waters, 1952–1983 and 1993–2004. Paper SC/58/RMP6 presented to the IWC Scientific Committee, May 2006, St. Kitts and Nevis, West Indies. 23pp. [Paper also available from the Office of this Journal].

Angliss, R. and Wade, P.R. 1997. Guidelines for assessing marine mammal stocks: report of the GAMMS Workshop, 3–5 April, 1996, Seattle, WA [Available at: www.nmfs.noaa.gov/pr/pdf/sars/gamms_report.pdf].

Butterworth, D.S. and Punt, A.E. 1994. An investigation of the merits or otherwise of the proposal for an Antarctic-wide whale sanctuary by means of adaptations of the simulation trials used to test the Revised Management Procedure. *Rep. Int. Whal. Commn.* 44: 289–302.

Chen, C., Durand, E., Forbes, F. and Francois, O. 2007. Bayesian clustering algorithms ascertaining spatial population structure: a new computer program and comparison study. *Mol. Ecol. Notes* 7(5): 747–56.

Gardenfors, U., Hilton-Taylor, C., Mace, G. and Rodriguez, J. 2001. The application of IUCN Red List criteria at region levels. *Cons. Biol.* 15: 1,206–12.

IWC. 1991. Report of the Scientific Committee, Annex D. Report of the sub-committee on management procedures. *Rep. int. Whal. Commn.* 41: 90–112C.

IWC. 1992a. Report of the Scientific Committee. *Rep. int. Whal. Commn.* 42: 51–86.

IWC. 1992b. Report of the Scientific Committee, Annex D. Report of the Sub-Committee on Management Procedures. *Rep. int. Whal. Commn* 42: 87–136.

IWC. 1994. Report of the Scientific Committee. Annex H. The Revised Management Procedure (RMP) for Baleen Whales. *Rep. int. Whal. Commn* 44: 145–52.

IWC. 2004. Report of the Scientific Committee. Annex D. Report of the Sub-Committee on the Revised Management Procedure. *J. Cetacean Res. Manage. (Suppl.)* 6: 75–184.

IWC. 2007. Report of the Scientific Committee. Annex D. Report of the Sub-Committee on the Revised Management Procedure. *J. Cetacean Res. Manage. (Suppl.)* 9: 88–128.

IWC. 2009. Report of the Scientific Committee. Annex D. Report of the Sub-Committee on the Revised Management Procedure (RMP). Appendix 7. Report of the working group on the North Atlantic minke whales RMP Implementation Review. *J. Cetacean Res. Manage. (Suppl.)* 11: 132–40.

Latch, E., Dharmarajan, G. and Glaubitz, J. 2006. Relative performance of Bayesian clustering software for inferring population substructure and individual assignment at low levels of population differentiation. *Conserv. Genet.* 7: 295–302.

Martien, K.K. and Taylor, B.L. 2003. Limitations of hypothesis-testing in defining management units for continuously distributed species. *J. Cetacean Res. Manage.* 5(3): 213–19.

Martien, K.K. 2006. Progress on TOSSM dataset generation. Paper SC/58/SD2 presented to the IWC Scientific Committee, May 2006, St Kitts and Nevis, West Indies (unpublished). 17pp. [Paper available from the Office of this Journal].

Martien, K.K., Gregovich, D., Bravington, M.V., Punt, A.E., Strand, A.E., Tallmon, D.A. and Taylor, B.L. 2009. TOSSM: an R package for assessing performance of genetic analytical methods in a management context. *Mol. Ecol. Res.* 9(6): 1,456–59.

Morin, P.A., Martien, K.K. and Taylor, B.L. 2008. Assessing statistical power of SNPs for population structure and conservation studies. *Mol. Ecol. Res.* 9: 66–73.

Moritz, C. 1994. Defining ‘evolutionary significant units’ for conservation. *Trends Ecol. Evol.* 9: 373–75.

Palsbøll, P.J., Berube, M. and Allendorf, F.W. 2007. Identification of management units using population genetic data. *Trends Ecol. Evol.* 22(1): 11–16.

Perryman, W.L., Donahue, M.A., Perkins, P.C. and Reilly, S.B. 2002. Gray whale calf production 1994–2000: Are observed fluctuations related to changes in seasonal ice cover? *Marine Mammal Science* 18: 121–24.

Punt, A.E. and Donovan, G. 2007. Developing management procedures that are robust to uncertainty: Lessons from the International Whaling Commission. *ICES J. Mar. Sci.* 64: 603–12.

Reilly, S.B. 1984. Observed and published rates of increase in gray whales, *Eschrichtius robustus*. *Rep. int. Whal. Commn.* (special issue) 6: 389–99.

Smith, T., Schweder, T., Allison, C., Palka, D., Punt, A.E. and Butterworth, D.S. In press. Effects of combinations of the risk performance of the IWC’s commercial whaling management procedure. *J. Cetacean Res. Manage.*

Strand, A.E. 2002. METASIM 1.0: an individual-based environment for simulating population genetics of complex population dynamics. *Mol. Ecol. Notes* 3: 373–76.

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

Taylor, B.L. 2005. Identifying units to conserve. pp.149–64. In: Reynolds III, J.E., Perrin, W.F., Reeves, R.R., Montgomery, S. and Regan, T.J. (eds). *Marine Mammal Research: Conservation Beyond Crisis*. The Johns Hopkins University Press, Baltimore, Maryland.

Taylor, B.L., Martien, K.K. and Morin, P.A. 2010. Identifying units to conserve using genetic data. In: Boyd, I., Bowen, D. and Iverson, S. (eds). *Marine Mammal Ecology and Conservation: a Handbook of Techniques*. Oxford University Press, Oxford, UK.

Taylor, B.L., Martinez, M., Gerrodette, T., Barlow, J. and Hrovat, Y.N. 2007. Lessons from monitoring trends in abundance of marine mammals. *Mar. Mammal Sci.* 23: 157–75.

Waples, R.S. 1991. Pacific salmon, *Oncorhynchus* spp., and the definition of ‘species’ under the Endangered Species Act. *Mar. Fish. Rev.* 53(3): 11–22.

Waples, R.S. and Gaggiotti, O. 2006. What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Mol. Ecol.* 15(6): 1419–39.

Appendix 1

TOSSM DATASET GENERATION

The TOSSM datasets used in this paper were generated following the methodology and parameter values described by Martien (2006). Rmetasim implements density dependence by linearly interpolating between matrices describing life history rates at zero population density (ZPD) and at carrying capacity (K)⁹. All 4% MSYR datasets were

generated using the life history matrices developed by Martien (2006; Table A1) from empirical data for Eastern Pacific grey whales (Perryman *et al.*, 2002; Reilly, 1984). The ZPD and K matrices developed by Martien (2006) have growth rates of $\lambda = 1.072$ and 1.0003, respectively, and result in an MSYR₁₊ of 3.9%.

Table A1

Life history matrices for used to generate the 4% MSYR datasets. Matrices describe life history parameters at: (a) zero population density; and (b) carrying capacity.

| (a) | juve1 | juve2 | fert | lact | male | (b) | juve1 | juve2 | fert | lact | male |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| juve1 | 0.730 | 0 | 0.889 | 0 | 0 | juve1 | 0.768 | 0 | 0.278 | 0 | 0 |
| juve2 | 0.210 | 0 | 0 | 0 | 0 | juve2 | 0.157 | 0.720 | 0 | 0 | 0 |
| fert | 0 | 0.47 | 0 | 0.946 | 0 | fert | 0 | 0.102 | 0.648 | 0.946 | 0 |
| lact | 0 | 0 | 0.946 | 0 | 0 | lact | 0 | 0 | 0.300 | 0 | 0 |
| male | 0 | 0.47 | 0 | 0 | 0.954 | male | 0 | 0.102 | 0 | 0 | 0.954 |

*Stage class abbreviations are juve1= juvenile1, juve2= juvenile2, fert=fertile female, lact=lactating female, and male=adult male.

⁹MSYL does not, therefore, occur at 0.6K as is conventional for most of the operating models used to test the CLA.

Generating datasets with an MSYR of 1% required developing a new *ZPD* matrix with a growth rate of $\lambda = 1.02$. We interpolated between the elements of the two matrices developed by Martien (2006; Table A1) to produce a matrix

with the desired growth rate (Table A2). When combined with the *K* matrix developed by Martien (2006), this matrix results in an $MSYR_{1+}$ of 1.0%.

Table A2

Life history matrices for used to generate the 1% MSYR datasets. Matrices describe life history parameters at: (a) zero population density; and (b) carrying capacity. The carrying capacity matrix is identical to that used to produce 4% MSYR datasets (Table A1b).

| (a) | juve1 | juve2 | Fert | lact | male | (b) | juve1 | juve2 | fert | lact | male |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| juve1 | 0.760 | 0 | 0.404 | 1.0 | 0 | juve1 | 0.768 | 0 | 0.278 | 0 | 0 |
| juve2 | 0.168 | 0.570 | 0 | 0 | 0 | juve2 | 0.157 | 0.720 | 0 | 0 | 0 |
| fert | 0 | 0.179 | 0.513 | 0.946 | 0 | fert | 0 | 0.102 | 0.648 | 0.946 | 0 |
| lact | 0 | 0 | 0.434 | 0 | 0 | lact | 0 | 0 | 0.300 | 0 | 0 |
| male | 0 | 0.179 | 0 | 0 | 0.954 | male | 0 | 0.102 | 0 | 0 | 0.954 |

Stage class abbreviations are juve1= juvenile1, juve2 = juvenile2, fert=fertile female, lact=lactating female, and male=adult male.