



# Improved abundance and trend estimates for sperm whales in the eastern North Pacific from Bayesian hierarchical modeling

Jeffrey E. Moore\*, Jay P. Barlow

Marine Mammal and Turtle Division, Southwest Fisheries Science Center, National Marine Fisheries Service, NOAA,  
8901 La Jolla Shores Drive, La Jolla, California 92037, USA

**ABSTRACT:** Population abundance and trends are informative metrics for assessing population status and basing management decisions, but it can be challenging to estimate these metrics for species that are difficult to detect. We used a Bayesian hierarchical approach to improve estimates of abundance and trends for sperm whales *Physeter macrocephalus* in the California Current based on 6 surveys conducted from 1991 to 2008. The method consists either of a regression trend or Markov process model for true abundance in the study area and an observation model used to estimate detection probability (in our case, based on line-transect distance sampling methods). Our approach improves upon models that we have applied to other cetacean species by incorporating direct estimates of sampling variance in the number of groups encountered (to reduce estimates of process variance) and by using an alternative to the commonly used Horvitz-Thompson-like estimator to better deal with covariate-dependent detection probability estimation when sample sizes are small. Our abundance estimates were much less variable through time than previously published estimates; this has important implications for management under the US Marine Mammal Protection Act. We were unable to precisely estimate overall abundance trends for sperm whales in the study area, but we estimated a high probability that small-group abundance, likely representing adult male abundance, has increased.

**KEY WORDS:** Detectability · Distance sampling · Marine Mammal Protection Act · Markov process · Overdispersion · *Physeter macrocephalus* · Population assessment

## INTRODUCTION

Estimates of population abundance and trends are among the most informative metrics upon which to assess a population's conservation status and base management decisions, but obtaining good inference from field survey data is often complicated by imperfect and heterogeneous detection of animals across surveys (e.g. Kéry et al. 2009). Methods abound for modeling detection probability to improve abundance estimation, but the added uncertainty contributed by the detection process nevertheless reduces precision of the abundance and trend estimates, often making it difficult to reliably assess population

status. For example, the US Marine Mammal Protection Act (MMPA) requires the National Marine Fisheries Service (NMFS) to routinely assess the population status of marine mammal stocks occurring regularly within US waters. Stock assessment reports (SARs) must include, among other things, estimates of population abundance and trends. For stocks inhabiting waters off the US West Coast (e.g. Carretta et al. 2013), abundance has been estimated from distance-sampling methods (Buckland et al. 2001, 2004), based on line-transect data from 6 ship-based survey cruises conducted between 1991 and 2008 (Barlow & Forney 2007, Barlow 2010). However, abundance trends have generally not been esti-

\*Corresponding author: jeff.e.moore@noaa.gov

mated, given the relatively few data points and typically low precision of individual estimates (e.g. coefficient of variation,  $CV > 0.3$  for most US stocks) that limit statistical power to detect cetacean population changes (Taylor et al. 2007, Jewell et al. 2012).

Bayesian hierarchical methods (e.g. Link & Sauer 2002, Royle & Dorazio 2008, Kéry & Royle 2010) can help improve inference about abundance trends and have recently been applied to US West Coast stocks of fin whales *Balaenoptera physalus* and beaked whales (Family Ziphiidae) (Moore & Barlow 2011, 2013). For these examples, a simple regression-trend model was used to describe change in population abundance through time as a function of a mean growth rate parameter with lognormal process variance to characterize annual fluctuations around the mean. An observation model (a detection model following line-transect distance sampling theory) described the relationship between modeled abundance and the field data. Trend and observation model parameters were estimated using Markov Chain Monte Carlo (MCMC) methods, and the estimates were expressed as posterior probability densities, allowing for probabilistic inference (e.g. probability that the population is declining or that abundance is above some value).

Here we applied an improved Bayesian hierarchical model to estimate population abundance and trends for sperm whales *Physeter macrocephalus* inhabiting waters of the US West Coast Exclusive Economic Zone (EEZ); these whales are designated by NMFS as the 'California-Oregon-Washington stock,' part of a larger but not clearly delineated North Pacific population (Mesnick et al. 2011, Mizroch & Rice 2013). Sperm whales are listed under the US Endangered Species Act as 'endangered' throughout their range, the species having been depleted by decades of whaling (Whitehead 2002, 2006). Adult males were especially targeted and thus disproportionately depleted (e.g. Whitehead et al. 1997). While many large whale stocks have been recovering following the 1986 global moratorium on commercial whaling, there is no evidence to date of sperm whale recovery in the North Pacific or elsewhere. This may simply be an artifact of insufficient data to assess trends for this difficult-to-survey species (Kaschner et al. 2012). It has also been hypothesized that severe disruption to the species' age- and complex societal structure may have had lingering consequences for population dynamics (Whitehead 2003, Wade et al. 2012).

Survey-specific estimates for the California-Oregon-Washington stock, obtained using distance sampling methods, have ranged from a high of 3140 ( $CV = 0.4$ )

in 2005 to a low of 300 ( $CV = 0.5$ ) in 2008 (Barlow & Forney 2007, Barlow 2010). Sperm whales are highly mobile and wide ranging, and our study area encompasses only a fraction of the entire population, so some real annual variation in abundance is expected within the study area, but high sampling variance is also an important contributor to variation in the stand-alone annual estimates. Sperm whale sightings are relatively rare (only 84 groups from the 6 surveys are included in our analysis, an average of  $14 \text{ yr}^{-1}$ ), and group size is highly variable (ranging in our dataset from 1 to 37), so a few chance observations can greatly affect individual-year estimates. Such high variation in annual abundance estimates can have important management implications, as we discuss later (see 'Discussion: Management implications').

The model we describe improves on our previous methods (Moore & Barlow 2011, 2013) in 2 respects. First, we input explicit estimates of sampling variance for the number of sperm whale groups sighted, which allows for improved (reduced) estimates of process variance and hence greater precision in trend estimates and projected abundance estimates. Second, effects of covariates on detection probability are commonly estimated using a Horvitz-Thompson-like estimator (Marques & Buckland 2004, Barlow & Forney 2007), which can be problematic in cases of small sample sizes. We use an alternative approach to circumvent this issue. Our analyses provide new estimates of sperm whale abundance that are higher and more stable across years than currently published values. Conclusive estimates of trends for total abundance within the study area were not obtained (low precision in the growth rate parameter), but we found that the abundance of adult male sperm whales has increased within the California Current. Our results have important and immediate implications for an ongoing management process for sperm whales in the USA.

## METHODS

### Data

Survey methods for these data have been described elsewhere in various detail (Kinzey et al. 2000, Barlow & Forney 2007, Moore & Barlow 2011). Briefly, data for the current analysis were collected during ship-based line-transect survey cruises conducted in late summer and early fall of 1991, 1993, 1996, 2001, 2005, and 2008 within the California Current marine ecosystem. These surveys were not de-

signed for sperm whales in particular but rather for collecting data on all observable cetacean species in the system. The 1.14 million km<sup>2</sup> study area extends to ca. 556 km (300 nautical miles) offshore the US West Coast, from the US–Canadian to US–Mexican border, stratified into 4 areas of similar size (Fig. 1). Stratification is important to analyses because effective survey effort can vary geographically (due to geographic variation in survey conditions). Total on-effort transect length within the study area was ca. 62 900 km (6200 to 14 700 km yr<sup>-1</sup>; see Moore & Barlow 2011 for tabulation by stratum within year). Observers on the ship detected cetacean groups visually using 25 × 150 binoculars mounted on the vessel’s flying bridge. For each sighting, group size (number of individuals) and perpendicular distance from the vessel were recorded. Environmental conditions that could potentially affect detection probability were regularly recorded along the effort transects (e.g. Beaufort sea state, rain or fog conditions, glare, estimated visibility). In total, 100 sperm whale groups were detected within 8 km of the vessel; however,

following recommendations by Buckland et al. (2001), it is common to truncate the data to eliminate the most distant observations, which are measured less accurately from the ship (using binocular reticles) and can make modeling the detection function difficult. Previous analyses of these data were based on a truncation distance of 4 km (Barlow & Forney 2007, Barlow 2010). Here (and following Barlow et al. 2011), we used a distance of 5.5 km to increase sample size; thus, 84 detected sperm whale groups were used in the analysis (Table 1).

### Analysis

#### Population trend model

We specified 2 different trend models, each describing simple exponential population growth. The first is a log-linear regression model:

$$N_t = N_0 \exp(rt) \exp(\epsilon_t) \tag{1}$$

$$\epsilon_t \sim \text{Normal}(0, \sigma^2_{\text{process}})$$

where annual abundance in year  $t$  ( $N_t$ ;  $t = 1, 3, \dots, 18$  for years 1991, 1993, ... 2008) is a function of an initial abundance parameter ( $N_0$ ), the mean per-capita growth rate ( $r$ ), and process error ( $\epsilon_t$ ), which we treat as log-normally distributed, although other error distributions, such as the gamma distribution, could be used. The growth rate  $r$  is the sum of 4 per-capita rates: birth – death + immigration – emigration. Each

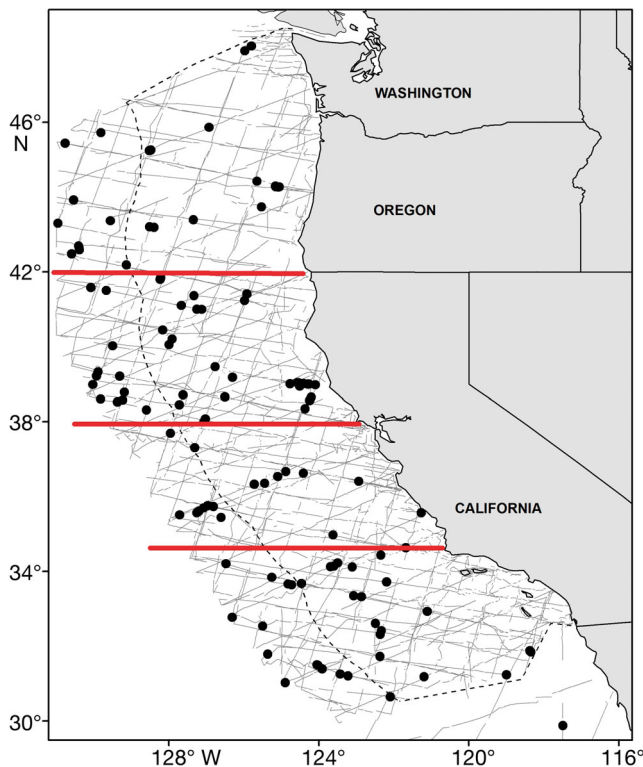


Fig. 1. Transects surveyed within the study area, divided into 4 geographic strata (from north to south: Oregon-Washington, Northern California, Central California, Southern California); US West Coast Exclusive Economic Zone (EEZ) boundary (dotted line); and sperm whale *Physeter macrocephalus* sighting locations (●)

Table 1. Number of sperm whale *Physeter macrocephalus* groups seen,  $n_{gjt}$ , by group size class ( $g = 1$  for small groups; 2 for large groups) within each stratum  $j$  and year  $t$ . The Oregon-Washington stratum was not surveyed in 1991 or 1993 (nd: no data)

Stratum	1991	1993	1996	2001	2005	2008
Oregon-Washington						
$n_{1jt}$	nd	nd	1	1	3	3
$n_{2jt}$	nd	nd	3	1	1	0
Northern California						
$n_{1jt}$	0	1	1	3	7	3
$n_{2jt}$	0	7	2	1	7	2
Central California						
$n_{1jt}$	0	0	1	1	0	3
$n_{2jt}$	2	4	2	0	0	0
Southern California						
$n_{1jt}$	4	0	2	0	2	1
$n_{2jt}$	7	1	1	4	1	1
Totals						
$n_{1t}$	4	1	5	5	12	10
$n_{2t}$	9	12	8	6	9	3

of these rates has an annual process error; we view  $\varepsilon_t$  as the sum of these, such that  $\sigma_{\text{process}}^2$  describes variation in abundance with the study area due to demographic and movement processes. The second model has the same parameters but describes population change as a Markov process:

$$\begin{aligned} N_1 &= N_0 \exp(r) \exp(\varepsilon_t) \\ N_t &= N_{t-1} \exp(r) \exp(\varepsilon_t) \\ \varepsilon_t &\sim \text{Normal}(0, \sigma_{\text{process}}^2) \end{aligned} \quad (2)$$

For sperm whales, the detection process differs for animals occurring singly or in pairs from those occurring in group sizes  $>2$  (see below). The former are presumed to be adult males, based on descriptions of social structure by Best (1979), Lettevall et al. (2002), Whitehead (2003) (also see 'Discussion'). Therefore, we divided estimates of total abundance into 2 group-size classes (denoted  $g = 1$  for small groups and  $g = 2$  for large groups), so that  $N_{gt} = N_t \varphi_{gt}$ . We treat  $\varphi_{1t}$  as a random variable:  $\text{logit}(\varphi_{1t}) \sim \text{Normal}(\mu_{\varphi 1}, \sigma_{\varphi 1})$ , and  $\varphi_{2t} = 1 - \varphi_{1t}$ . The  $N_{gt}$  were further partitioned across the 4 survey strata ( $j$ ):

$$\begin{aligned} N_{gjt} &= N_{gt} \varphi_{gjt} \\ \varphi_{gjt} &\sim \text{Dirichlet}(u_{gj}) \end{aligned} \quad (3)$$

The Dirichlet distribution is a multivariate generalization of the beta distribution. For a given group-size class, the Dirichlet parameters  $u_{g1}, u_{g2}, u_{g3}, u_{g4}$  ( $u > 0$ ) dictate the mean proportions in each stratum and sum to 1. Variance in  $\varphi_{gjt}$  decreases as the sum of  $u_{gj}$  increases, so higher total  $u_{gj}$  would imply greater consistency across years in how abundance is distributed across the survey strata. The density ( $D$ ) of animals belonging to each group-size class, by stratum and year, is  $D_{gjt} = N_{gjt} / A_j$ , where  $A_j$  is the stratum area.

#### Observation model

The primary data are the number of groups belonging to group-size class  $g$  detected within each survey stratum  $j$  during survey year  $t$ , i.e.  $n_{gjt}$ . In our earlier analyses (Moore & Barlow 2011, 2013), we assumed that  $n$  was Poisson distributed; this may have underestimated sampling variance and thus overestimated process variance. In some cases, Poisson overdispersion may be handled simply by adding an extra random effect term to the observation model (e.g. Link & Sauer 2002, Kéry et al. 2009), but estimates of extra-Poisson sampling variance and process variance can be confounded, particularly in the case of small samples or if there is a lack of replication within the ran-

dom-effect level. We dealt with overdispersion by treating  $n$  as a random variable from a generalized Poisson process (Famoye 1993, Famoye et al. 2004):

$$n_{gjt} \sim \text{gPoisson}(\mu_{gjt}, \alpha_{gjt}) \quad (4)$$

where  $\alpha_{gjt}$  is an overdispersion parameter. We used this error distribution because it conveniently allowed us to incorporate independent empirical estimates of extra-Poisson sampling variance into the model framework. Based on a bootstrap approach in which we repeatedly calculated  $n_{gjt}$  from re-samples of survey line-segments, we estimated the average variance-mean ratio for  $n_{gjt}$  to be 1.3 for small groups and 2.3 for large groups (in contrast with a value of 1 for a simple Poisson process); it was straightforward to then estimate  $\alpha_{gjt}$  within the generalized Poisson likelihood (see Supplement 1 at [www.int-res.com/articles/suppl/n025p141\\_supp.pdf](http://www.int-res.com/articles/suppl/n025p141_supp.pdf)).

The estimator for the expected number of groups detected,  $\mu_{gjt}$ , has a basic form:

$$\mu = \frac{D}{s} \times 2LW \times p \quad (5)$$

where  $D$  is animal density,  $s$  is mean group size (so  $D/s$  is the group density),  $L$  is the total transect length,  $W$  is the maximum (truncation) distance from the vessel within which groups are detected (so that  $2LW$  is the total area surveyed;  $W = 5.5$  km in our analysis), and  $p$  is the average detection probability for groups occurring within the area surveyed. For

line-transect sampling,  $p = \frac{1}{W} \times \frac{g(0)}{f(0)}$ , where  $g(0)$  is detection probability of groups on the trackline, and  $1/f(0)$  is the so-called effective strip width, with  $f(0)$  being the value at distance  $y = 0$  of the probability density function for the detection function  $g(y)$  (Buckland et al. 2001). If detection probability varies as a function of covariates ( $\mathbf{z}$ ), then it is common to use a Horvitz-Thompson-like estimator, for which the estimate of  $p$  is conditioned on the sightings (Marques & Buckland 2004, Barlow & Forney 2007), so that if  $p$  is independent of group size:

$$\mu = \frac{D}{s} \times 2L \times \frac{n}{\sum_{i=1}^n f_i(0 | \mathbf{z}_i) / g_i(0 | \mathbf{z}_i)} \quad (6)$$

or if  $p$  depends on group size (e.g. larger groups are more detectable), then:

$$\mu = D \times 2L \times \frac{n}{\sum_{i=1}^n s_i f_i(0 | \mathbf{z}_i) / g_i(0 | \mathbf{z}_i)} \quad (7)$$

These estimators (see Supplement 2 at [www.int-res.com/articles/suppl/n025p141\\_supp.pdf](http://www.int-res.com/articles/suppl/n025p141_supp.pdf) for sim-

ple derivation) are asymptotically unbiased but can be problematic when sample sizes are small such that estimates of  $p$  or  $p/s$  in a particular stratum are calculated from just a few groups. Even more problematic, as would be the case in our analysis, is when some of the  $n_{gjt} = 0$ , so that  $\mu_{gjt} = 0$ , and the lack of detections is treated as no information in the likelihood. Better would be to estimate  $p$  solely as a function of survey conditions, rather than conditioning on the sightings.

If the line-transect effort is divided into  $K$  segments (or, for example, goes through habitat partitioned into a gridded map with  $K$  cells), with detection-related covariates measured for each, and detection probability also varies with group size, then a general estimator is:

$$\mu = \sum_{s=1}^s \sum_{k=1}^K \mu_{sk} \tag{8}$$

$$\mu_{sk} = \frac{D_s}{s} \times 2L_k \times \frac{g_k(0 | \mathbf{z}_k, s)}{f_k(0 | \mathbf{z}_k, s)} \tag{9}$$

Note that  $D_s$  (density of animals occurring in groups of size  $s$ ) and  $s$  could also be modeled for each  $k$  as a function of covariates (e.g. Royle et al. 2004, Gerrodette & Eguchi 2011), but here they are represented as mean values (e.g. for the year or stratum encompassing the  $K$  effort segments). Modeling group-size specific  $\mu$  at the segment-level may be impractical. However, the above estimator can be variously simplified for individual cases, for example, if group sizes can be collapsed into just a few classes ( $g$ ) within which  $p$  is constant (e.g. small and large group sizes, with a mean  $s$  for each), or if  $p$  is inde-

pendent of group size entirely (so that summing across group sizes or size classes is unnecessary, and a single mean  $s$  could be used). The estimator could also be simplified if  $p$  is modeled as a function of just 1 or a couple categorical covariates. We illustrate such a simplification with our analysis.

For our sperm whale dataset, the detection process differed for 2 group-size classes (Fig. 2). For animals occurring singly or in pairs ( $n = 37$  groups), detectability decreased with distance from the vessel. For larger groups ( $n = 47$ ), there was no apparent decrease in detection probability from distance 0 to  $W$ , so that for large groups, we were essentially conducting a strip transect (animals at all distances detected with same probability,  $g(0)$ ). We therefore assumed a value of  $f(0) = 1/W$  for large groups (i.e. effective strip width =  $W$ ). We estimated  $f(0)$  for small groups using a half-normal key function (Buckland et al. 2001) and an intercept-only model fit to the distance data, as the small sample size precluded covariate model fitting (MCMC chains would not converge). However,  $g(0)$ , and therefore  $p$ , decreases as Beaufort sea state increases (J. P. Barlow unpubl. data; Table 2). It may also differ by group size class, but this has not been estimated, so we modeled  $g(0)$  the same for both size classes. By specifying  $g(0)$  as a function of a single categorical covariate (sea state), we could collapse individual line-segments into the total length of effort occurring in each level of the covariate. Average  $g(0)$  was therefore weighted by the proportion of transect effort occurring in different sea states but was assumed to be the same for both group size classes. Our simplified estimator for  $\mu_{gjt}$  was:

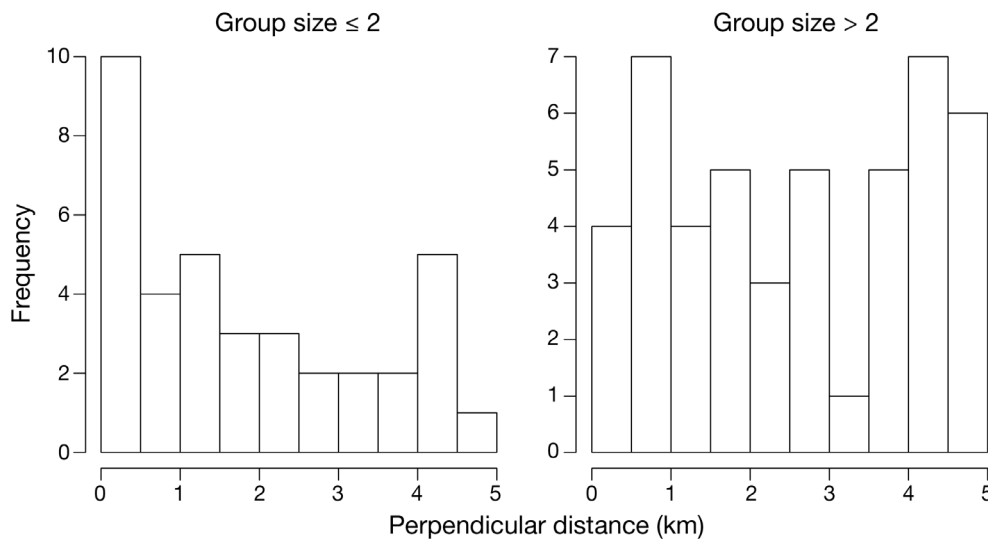


Fig. 2. Perpendicular detection distances, within a truncation distance of 5.5 km of the vessel, for small and larger sperm whale *Physeter macrocephalus* groups (no groups were detected between 5 and 5.5 km)

$$\mu_{gjt} = \frac{D_{gjt}}{s_{gjt}} \times 2L_{jt} \times \frac{\sum_{b=1}^6 L_{jtb} g_b(0) / f_g(0)}{\sum_{b=1}^6 L_{jtb}} \quad (10)$$

where  $b$  are the 6 Beaufort sea state classes: 0, 1, ..., 5. Note that in previous analyses (e.g. Barlow & Forney 2007, Barlow 2010, Whitehead 2002), a value of 0.87 (CV = 0.09) was used for  $g(0)$  in all sea states. Thus, in our analysis,  $g(0)$  is lower overall and more uncertain than has been assumed for previous analyses.

For group size, we fixed  $s_{1jt} = 1.14$ , the mean group size of small groups (again,  $g = 1$ ) in the dataset. Most small-group sightings consisted of single individuals, and we considered the small standard error for the observations (0.05) as negligible. For large groups ( $g = 2$ ), we did not want to rely solely on the observed sizes in each  $j, t$  to estimate  $s_{2jt}$  because of small samples within these individual strata. Alternatively, using the grand mean for all  $j, t$  would ignore possible real variation across years and geographic strata. We therefore treated  $s_{2jt}$  as a random effect variable, generating estimates of  $s_{2jt}$  that are intermediate between a grand mean and the individual data means for each  $j, t$  (Supplement 3 at [www.int-res.com/articles/suppl/n025p141\\_supp.pdf](http://www.int-res.com/articles/suppl/n025p141_supp.pdf)). Also, less effort was spent estimating sperm whale group size during the first 3 surveys (in the 1990s) than in later surveys (in the 2000s); large-group sizes for the early years are known to be underestimated for this reason. We therefore estimated an extra parameter in the model as a correction factor. See Supplement 3 for model details for  $s_{2jt}$ .

#### Parameter estimation

Analyses were conducted in OpenBUGS 3.2.2, called by R using the library R2OpenBUGS. Posterior distributions for estimated parameters were constructed from 2 MCMC chains, each of length 500 000. The first 100 000 were discarded (burn-in), and the remaining samples were thinned by 10, so that posterior distributions were comprised of 80 000 retained samples.  $\hat{R}$  values (an MCMC convergence statistic outputted by OpenBUGS) of 1.00 on all parameters suggested satisfactory model convergence. Flat priors were used for all model parameters to be estimated (see Table S1 in Supplement 4 and Supplement 5 at [www.int-res.com/articles/suppl/n025p141\\_supp/](http://www.int-res.com/articles/suppl/n025p141_supp/)). Post hoc analysis of trends in adult male abundance Posterior estimates for the  $N_{1t}$  (abundance of individuals occurring as loners or pairs, presumed to be adult males) generally increased with time. To evaluate this increase statistically, we fit a simple log-linear trend

Table 2. Bayesian priors for probability of detecting a sperm whale *Physeter macrocephalus* group on the vessel transect line (perpendicular distance from ship = 0 km),  $g_b(0)$ , for Beaufort sea state  $b$ . Values for Beaufort 0 and 1 conditions,  $g_1(0)$ , are from Barlow & Taylor (2005) and Barlow & Forney (2007). Estimates for higher sea states ( $b = 2, 3, 4, 5$ ) were calculated as  $g_0(0)e^{-0.113b}$ , with variance on the slope coefficient chosen to provide estimates and CVs approximating those by Barlow (in review)

Beaufort sea state	Mean $g(0)$	SE
0	0.87	0.08
1	0.87	0.08
2	0.69	0.14
3	0.63	0.17
4	0.57	0.20
5	0.52	0.22

model to each posterior sample of  $N_{1t}$  estimates and summarized the posterior distribution of trend model slope coefficients.

## RESULTS

### Detection probability, population trends, and abundance

Inferences from regression and Markov trend models were almost identical, with a couple of exceptions noted below. Unless stated otherwise, values reported in this section are from the regression model, for simplicity. We generally report posterior medians for parameter estimates, but we report posterior modes for abundance estimates because the distributions are strongly right-skewed and the mode (the most probable value) is perhaps the most comparable to the maximum-likelihood point estimates that have been reported previously (note, however, that Table 3 includes the posterior mode, median, and mean for

Table 3. Posterior summaries for total sperm whale *Physeter macrocephalus* abundance, based on the regression trend model. Markov process model estimates were almost identical, i.e. point estimates and CVs were within a few percent of these values

Year	20 <sup>th</sup> percentile	Mode	Median	Mean	CV
1991	1249	1395	2011	2388	0.64
1993	1411	1611	2258	2729	0.67
1996	1341	1569	2024	2303	0.54
2001	1445	1750	2160	2439	0.52
2005	1722	2081	2628	3013	0.55
2008	1332	1658	2106	2431	0.58

abundance). The effective strip width,  $1/f(0)$ , for small groups (group size  $\leq 2$ ) was estimated to be 3.4 km (95% credible interval, CRI: 2.6–5.2 km), compared to the fixed value of 5.5 km used for large groups. Thus if  $g(0)$  were 1, detection probability  $p$  within a 5.5 km half-strip would be 0.62 for small groups (individuals and pairs) and 1.00 for large groups. However,  $p$  depends on both effective strip width and Beaufort-dependent  $g(0)$ , which varies by survey year according to Beaufort sea state conditions. We estimate that for small groups,  $p$  within the surveyed area ( $2LW$ ) ranged from 0.31 to 0.36 across different strata and years, while  $p$  for larger groups ranged from 0.50 to 0.58. However, substantial uncertainty in  $g(0)$ , particularly in higher Beaufort states where a majority of survey effort occurs (Table 2), contributed to low precision in the  $p_{gjt}$  (CVs  $\approx 0.24$ –0.38).

For the first 3 survey years, corrected mean group size for large groups (group size  $> 2$ ) was ca.  $2.3\times$  higher on average than uncorrected values (95% CRI:  $1.1$ – $5.3\times$ ). Mean corrected group size for large groups was 16 (95% CRI: 9–29) overall, with corrected model estimates ranging from 8 to 24 across all individual years and strata.

Annual rate of change,  $r$ , for total sperm whale abundance was 0.008 and 0.006 for the regression and Markov models, respectively, but precision was low; 95% CRIs ranged from  $-0.09$  to  $+0.10$  for the regression model and from  $-0.17$  to  $+0.18$  for the Markov model. A simpler regression model for which group size was kept constant across all  $j$ ,  $t$  (i.e. random effect term omitted)—this simply evaluates trend in number of groups detected—provided almost identical results for  $r$  (not shown). The probability that  $r$  was positive was 0.57 and 0.54, an equivocal result. For the regression model,  $\sigma_{\text{process}} = 0.27$ , whereas for the Markov model,  $\sigma_{\text{process}} = 0.19$ . Because sampling variance in  $n_{gjt}$  was greater than Poisson, our inclusion of explicit sampling variance estimates helped reduce the estimates of  $\sigma_{\text{process}}$ , which would have otherwise included both true process and extra-Poisson sampling variance. This improves precision in the trend parameter and allows for more precise prediction and projection modeling.

Nevertheless, these are fairly large process variance estimates, indicative of an open population if they are accurate. For example, the regression model suggests that increases in annual abundance of  $>30\%$  or decreases of  $>23\%$  from one year to the next occur in more than half of all years, with occasional doubling or halving (once in 10 yr). With maximum potential population growth for sperm whales probably being less than a few percent per year (Whitehead 2002), such population variation would not be possible from birth and death processes alone.

Posterior modes for abundance (Table 3, Fig. 3A) ranged from a low of 1395 in 1991 to a high of 2081 in 2005; CVs for individual year estimates ranged from 0.52 to 0.67. Of particular relevance for the sperm whale SAR prepared by NMFS, the 20<sup>th</sup> percentile

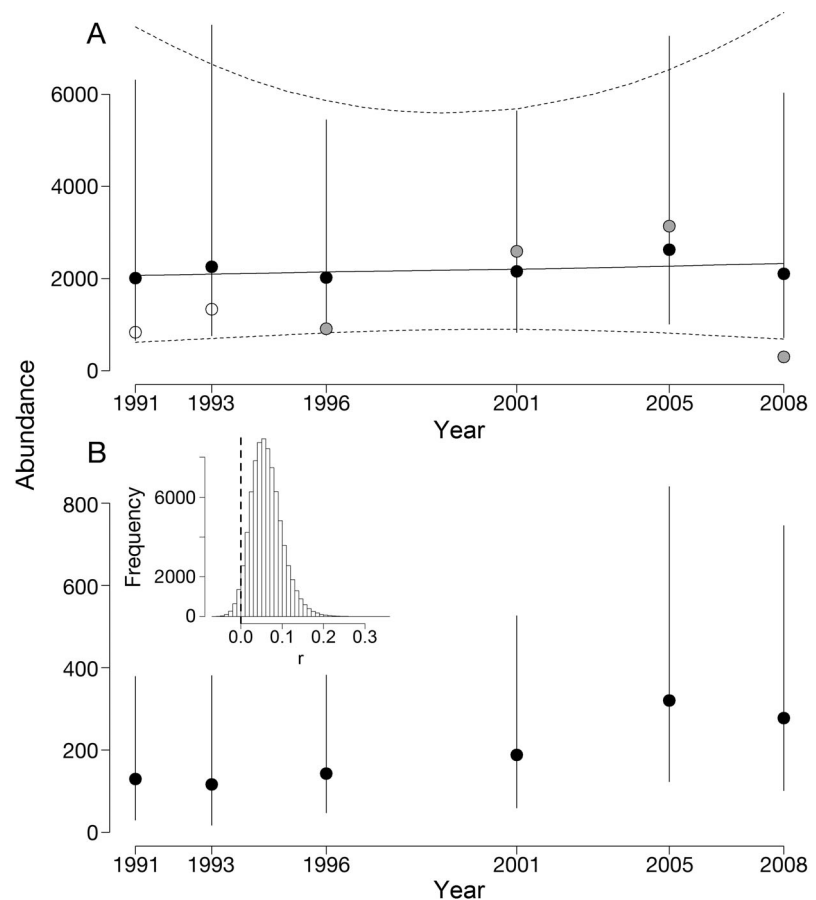


Fig. 3. (A) Abundance estimates (posterior medians [black circles] and 95% credible intervals, CRIs) from the regression trend model for sperm whales *Physeter macrocephalus*, with fitted trend line and 95% CRIs for trend. For comparison, open and gray circles depict earlier published estimates from Barlow & Forney (2007) and Barlow (2010), with those for 1991 and 1993 (open circles) being for a smaller surveyed area. (B) Abundance estimates (black circles, with 95% CRIs) and posterior distribution (inset) for annual rate of change for small groups (individuals occurring as loners or pairs, presumably adult males). In both panels, abundance estimates from the Markov model are virtually identical to the regression model estimates

estimate of abundance ( $N_{\min}$ ) for 2008 (corresponding to the most recent survey year) was 1332 for the regression model and 1358 for the Markov model, compared to the currently used value of 751.

### Increasing trend for abundance of adult males

There was strong evidence that abundance of sperm whales occurring as lone individuals or in pairs (very likely adult males) has increased in the study area. For both the regression and Markov model, the posterior median and mean growth rate was 0.06 (SD = 0.038), and the probability of the growth rate being  $>0$  was 0.97. Abundance for this segment of the population more than doubled over the course of the survey period, with posterior median estimates ranging from 117 (SD = 97) in 1993 to 321 (SD = 188) in 2005 (Fig. 3B).

## DISCUSSION

### Comparison to previous estimates

Compared to previous abundance estimates for the California Current study area, our estimates were higher and less variable. The geometric mean of our estimates across all 6 survey years was 1665 (posterior modes), 2189 (medians), or 2539 (means), compared to a previous point estimate of 1529 for data pooled across these years (Barlow 2010). Our point estimates (to use modes) ranged from ca. 1400 to 2100, compared to previous estimates ranging from 300 to 3140. The overall increase is expected because we used lower  $g(0)$  values and corrected group-size underestimates for the earlier surveys. Other factors would have had less predictable effects, such as our inclusion of additional data (by extending data truncation distance from 4 km in previous analyses to 5.5 km in the current one) and differences in the model approach. Reduced variation in the estimates was the result of including a trend model that effectively estimates annual abundance from fewer parameters, using shrinkage estimators by treating certain parameters (e.g. group size in  $j, t$ ) as random effect variables, and partitioning the detection process by group-size class. These aspects of the model also increase the precision of estimates and improve the ability to detect trends. Our annual estimates were actually less precise than previous estimates, but this is only because we included additional forms of uncertainty in our analysis here, namely higher

CVs for sea state-specific estimates of  $g(0)$ , and estimation (with error) of a group size correction factor.

### Management implications

In any management context, high variation in annual abundance estimates can have important implications. Our case study exemplifies the benefit of reducing this variation. The MMPA requires for each marine mammal stock an estimate of potential biological removal (PBR) – an annual limit to mortality and serious injury tied to population goals under the Act, calculated as  $0.5R_{\max}N_{\min}F_R$ , where  $R_{\max}$  is maximum potential growth rate,  $N_{\min}$  is the 20th percentile estimate of abundance, and  $F_R$  is a recovery factor that varies according to management considerations. Thus, variation in PBR follows variation in the abundance estimates. If PBR is exceeded, a Take-Reduction Plan (TRP) process is initiated to reduce fisheries-related impact. This process is costly to government and fisheries, so it is desirable to initiate only when necessary. For sperm whales, the most recent abundance point estimate (from the 2008 survey) is only 300, which may be inaccurate given high sampling variance. To buffer this volatility, the most recent SAR for sperm whales uses a geometric mean from the 2005 and 2008 estimates (Carretta et al. 2013), but this remains less than ideal because it still only makes use of some information from the time series, and it ignores any potential population trends that may be occurring (e.g. a simple mean from multiple years is an inherently biased estimate of current abundance if the population is not stable). Moreover, the current SAR guidelines (NMFS 2005) recommend that abundance estimates used in the assessments be  $<8$  yr old. Therefore, the 2005 estimate has now 'expired,' so if the sperm whale SAR were revised today, the abundance estimate would substantially decrease. The PBR estimate from the most recent SAR is 1.5. A new PBR value based on the 2008-only abundance estimate would be  $<1$ . The results we have presented translate to PBR = 2.7. These differences are non-trivial, because, depending on how many years of data are pooled, the average estimate for annual fisheries mortality and serious injuries is around 1.3 to 1.6 over the past 10 to 12 yr of data (2001–2012), and a Take-Reduction Team is currently revising its TRP (which involves new fishery regulations and revisions to the scientific observer program) to deal with sperm whale bycatch issues. Clearly there are ramifications of how abundance is estimated.



### Benefits of the hierarchical Bayesian approach

Our analysis demonstrates several benefits of taking a hierarchical Bayesian approach to trend estimation. Posterior distribution summaries for parameters facilitate probabilistic inference, making it straightforward to directly estimate the probability of a positive or negative trend. Model specification is highly flexible (including Markov processes) and missing data are handled well. In our case, it was straightforward to partition the modeling process for different group-size classes, specify generalized Poisson error distributions with empirical inputs for variance inflation and overdispersion parameters, and deal with missing sightings or group size information from some of the  $g,j,t$  combinations.

Improved inference from incorporating trend models into the abundance estimation can be explained in model selection terms. In previous sperm whale analyses, data were pooled across years to estimate the detection function, but group size and abundance estimates for each stratum  $\times$  year were otherwise obtained from data unique to those cells. This is like fitting a fully stratum  $\times$  time-dependent abundance model, requiring a separate parameter for group size and abundance in each cell, with relatively few data points per parameter and hence low precision. In contrast, the trend model formalizes an assumption that abundance within the survey area is related through time, movements in and out of the study area notwithstanding; thus, the population estimate for any one year is informed by the entire dataset rather than just that year's data. The estimates in this case are obtained from fewer parameters (an intercept, a slope coefficient, and a random-effect process-variance estimate), with more data per parameter and thus increased precision. Treating true abundance as a random variable has the effect of limiting the extent to which individual abundance estimates can vary, increasing their precision and shrinking stand-alone estimates toward a modeled expectation. In general, simpler models (fewer parameters) improve precision but at the possible expense of increased bias (Burnham & Anderson 2002). Formal model selection metrics (e.g. deviance information criterion, DIC; Spiegelhalter et al. 2002) can help guide the search for parsimony, but model selection criteria for random-effect models are generally lacking and not always necessary. We opted simply to fit the most general model that made biological sense on *a priori* grounds and for which enough data were available to provide satisfactorily precise and reasonable estimates.

### Inference about sperm whales in the eastern North Pacific

We were unable to obtain good estimates of abundance trends for the entire California-Oregon-Washington stock of sperm whales. Our best estimate is that numbers in the California Current study area were stable from 1991 to 2008, but precision of the growth rate estimate is too low to make any reliable conclusion. However, while a trend in total numbers is difficult to assess, there is a high probability that the number of animals traveling alone (or in some cases in pairs) has increased within the study area, the best estimate being more than a 2-fold increase from 1991 to 2008. These are likely to be reproductive adult males, known to roam individually or in very small groups, whereas females and young animals occur in larger familial groups, or in small bachelor herds in the case of younger pre-breeding adult males (Best 1979, Lettevall et al. 2002, Whitehead 2003). Whether this trend reflects a population-level increase in adult male abundance or merely increased use of the study area by adult males is not possible to say from the data. If it is a population-wide phenomenon, this does not necessarily indicate that population numbers as a whole increased throughout the survey period; it could mainly reflect an ageing population, as male calves from the 1980s (at cessation of the large-scale industrial whaling era) would have been recruiting into the reproductive adult class during the 2000s. However, there are major uncertainties in sperm whale stock structure throughout the eastern North Pacific. Males and females may seasonally occupy different habitats, with females generally occurring in warmer and lower-latitude waters than males, and males known to range over great distances. Thus, it is unknown as to which demographic group the adult males in our dataset belong.

It has been hypothesized that a deficit of adult males in the population due to targeted depletion by industrial whalers may have stymied sperm whale population recovery to date (Whitehead et al. 1997). Increasing male abundance might be a more precisely measurable indicator of increased population numbers overall, or it might at least indicate an increase in the availability of suitable reproductive males. Future survey data and analyses are needed to estimate changes in sperm whale population growth throughout a larger study region (e.g. combining datasets summarized by Kaschner et al. 2012) and to evaluate other demographic indicators of population growth, such as the number of calves (or number of groups with calves) in the dataset.

*Acknowledgements.* We thank the marine mammal observers, cruise leaders, survey coordinators, officers, and crew who worked long hours in difficult conditions to collect the data used in this report. We are grateful to Barb Taylor, Tim Gerrodette, and 3 anonymous reviewers for their comments to improve the manuscript.

## LITERATURE CITED

- Barlow J (2010) Cetacean abundance in the California Current estimated from a 2008 ship-based line-transect survey. NOAA Tech Memo NOAA-TM-NMFS-SWFSC-456. NOAA, La Jolla, CA
- Barlow J, Forney KA (2007) Abundance and the population density of cetaceans in the California Current ecosystem. Fish Bull 105:509–526
- Barlow J, Tayler BL (2005) Estimates of sperm whale abundance in the northeastern temperate Pacific from a combined acoustic and visual survey. Mar Mamm Sci 21: 429–445
- Barlow J, Ballance LT, Forney K (2011) Effective strip widths for ship-based line-transect surveys of cetaceans. NOAA Tech Memo NOAA-TM-NMFS-SWFSC-484. NOAA, La Jolla, CA
- Best PB (1979) Social organization in sperm whales, *Physeter macrocephalus*. In: Winn HE, Olla BL (eds) Behavior in marine animals. Plenum, New York, NY, p 227–289
- Buckland ST, Anderson DR, Burnham KP, Laake JL, Borchers DL (2001) Introduction to distance sampling. Oxford University Press, New York, NY
- Buckland ST, Anderson DR, Burnham KP, Laake JL, Borchers DL, Thomas L (eds) (2004) Advanced distance sampling. Oxford University Press, New York, NY
- Burnham KP, Anderson DR (2002) Model selection and multimodel inference, 2nd edn. Springer, New York, NY
- Carretta JV, Oleson E, Weller DW, Lang AR and others (2013) U.S. Pacific Marine Mammal Stock Assessments: 2012. NOAA Tech Memo NOAA-TM-NMFS-SWFSC-504. US Department of Commerce, NOAA, La Jolla, CA
- Famoye F (1993) Restricted generalized Poisson regression model. Comm Stat Theory Methods 22:1335–1354
- Famoye F, Wulu JT Jr, Singh KP (2004) On the generalized Poisson regression model with an application to accident data. J Data Sci 2:287–295
- Gerrodette T, Eguchi T (2011) Precautionary design of a marine protected area based on a habitat model. Endang Species Res 15:159–166
- Jewell R, Thomas L, Harris CM, Kaschner K, Wiff R, Hammond PS, Quick NJ (2012) Global analysis of cetacean line-transect surveys: detecting trends in cetacean density. Mar Ecol Prog Ser 453:227–240
- Kaschner K, Quick NJ, Jewell R, Williams R, Harris CM (2012) Global coverage of cetacean line-transect surveys: status quo, data gaps and future challenges. PLoS ONE 7:e44075
- Kéry M, Royle JA (2010) Hierarchical modelling and estimation of abundance and population trends in metapopulation designs. J Anim Ecol 79:453–461
- Kéry M, Dorazio RM, Soldaat L, van Strien A, Zuiderwijk A, Royle JA (2009) Trend estimation in populations with imperfect detection. J Appl Ecol 46:1163–1172
- Kinzey D, Olson P, Gerrodette T (2000) Marine mammal data collection procedures on research ship line-transect surveys by the Southwest Fisheries Science Center. SWFSC Administrative Report LJ-00-08. NOAA, La Jolla, CA
- Lettevall E, Richter C, Jaquet N, Slooten E and others (2002) Social structure and residency in aggregations of male sperm whales. Can J Zool 80:1189–1196
- Link WA, Sauer JR (2002) A hierarchical analysis of population change with application to cerulean warblers. Ecology 83:2832–2840
- Marques FFC, Buckland ST (2004) Covariate models for the detection function. In: Buckland ST, Anderson DR, Burnham KP, Laake JL, Borchers DL, Thomas L (eds) Advanced distance sampling. Oxford University Press, New York, NY, p 31–47
- Mesnick SL, Taylor BL, Archer FI, Martien KK and others (2011) Sperm whale population structure in the eastern and central North Pacific inferred by the use of single-nucleotide polymorphisms, microsatellites and mitochondrial DNA. Mol Ecol Resour 11(Suppl s1):278–298
- Mizroch SA, Rice DW (2013) Ocean nomads: distribution and movements of sperm whales in the North Pacific shown by whaling data and Discovery marks. Mar Mamm Sci 29:E136–E165
- Moore JE, Barlow JP (2011) Bayesian hierarchical estimation of fin whale abundance trends from a 1991–2008 time series of line-transect surveys in the California Current. J Appl Ecol 48:1195–1205
- Moore JE, Barlow JP (2013) Declining abundance of beaked whales (Family Ziphiidae) in the California Current Large Marine Ecosystem. PLoS ONE 8:e52770
- NMFS (National Marine Fisheries Service) (2005) Revisions to guidelines for assessing marine mammals stocks. Available at [www.nmfs.noaa.gov/pr/pdfs/sars/gamms2005.pdf](http://www.nmfs.noaa.gov/pr/pdfs/sars/gamms2005.pdf) (accessed on 8 September 2014)
- Royle JA, Dorazio RM (2008) Hierarchical modeling and inference in ecology. Academic Press, London
- Royle JA, Dawson DK, Bates S (2004) Modeling abundance effects in distance sampling. Ecology 85:1591–1597
- Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A (2002) Bayesian measures of model complexity and fit. J R Stat Soc B 64:583–639
- Taylor BL, Martinez M, Gerrodette T, Barlow J, Hrovat YN (2007) Lessons from monitoring trends in abundance of marine mammals. Mar Mamm Sci 23:157–175
- Wade PR, Reeves RR, Mesnick SL (2012) Social and behavioural factors in cetacean responses to overexploitation: Are odontocetes less ‘resilient’ than mysticetes? J Mar Biol 2012:567276, <http://dx.doi.org/10.1155/2012/567276>
- Whitehead H (2002) Estimates of the current global population size and historical trajectory for sperm whales. Mar Ecol Prog Ser 242:295–304
- Whitehead H (2003) Sperm whales: social evolution in the ocean. University of Chicago Press, Chicago, IL
- Whitehead H (2006) Sperm whales in ocean ecosystems. In: Estes JA, DeMaster DP, Doak DF, Williams TM, Brownell RL Jr (eds) Whales, whaling, and ocean ecosystems. University of California Press, Berkeley, CA, p 324–334
- Whitehead H, Christal J, Dufault S (1997) Past and distance whaling and the rapid decline of sperm whales off the Galapagos Islands. Conserv Biol 11:1387–1396