Biological Conservation 142 (2009) 1542–1547

Contents lists available at ScienceDirect

Biological Conservation

journal homepage: www.elsevier.com/locate/biocon

Short communication

Cost-effective abundance estimation of rare animals: Testing performance of small-boat surveys for killer whales in British Columbia

Rob Williams [a,b,*], Len Thomas [c]

[a] Sea Mammal Research Unit, Gatty Marine Laboratory, University of St. Andrews, St. Andrews, Fife KY16 8LB, Scotland, UK
[b] Marine Mammal Research Unit, University of British Columbia, 2202 Main Mall, Room 247, AERL, Vancouver, BC, Canada V6T 1Z4
[c] Research Unit for Wildlife Population Assessment, Centre for Research into Ecological and Environmental Modelling, University of St Andrews, Fife KY16 9LZ, Scotland, UK

ARTICLE INFO

Article history:
Received 8 July 2008
Received in revised form 10 December 2008
Accepted 30 December 2008
Available online 13 March 2009

Keywords:
Abundance
Cetacean
Distance sampling
Line transect
Mark-recapture
Model averaging
Small-boat survey

ABSTRACT

Top predators are often rare, subject to anthropogenic mortality, and possess life-history traits that make them inherently vulnerable to extinction. IUCN criteria recognise populations as Critically Endangered when abundance is <250 mature individuals, but estimating abundance of rare species can be more challenging than for common ones. Cost-effective methods are needed to provide robust abundance estimates. In marine environments, small boats are more widely accessible than large ships for researchers conducting sightings surveys with limited funds, but studies are needed into efficacy of small-boat surveys. This study compares line transect and mark-recapture estimates from small-boat surveys in summer 2004 and 2005 for 'northern resident' killer whales in British Columbia to true population size, known from censuses conducted by Fisheries and Oceans Canada. The line transect estimate of 195 animals (95% CI 27–559) used model averaging to incorporate uncertainty in the detection function, while the mark-recapture estimate of 239 animals (CI 154–370) used a simple two-sample Chapman estimator. Both methods produced estimates close to the true population size, which numbered 219 animals in 2004 and 235 in 2006, but both suffered from the small sample sizes and violations of some model assumptions that will vex most pilot studies of rare species. Initial abundance estimates from relatively low-cost surveys can be thought of as hypotheses to be tested as new data are collected. For species of conservation concern, any cost-effective attempt to estimate absolute abundance will assist status assessments, as long as estimates are presented with appropriate caveats.

1. Introduction

Generating robust estimates of the abundance of rare animals is perennially difficult, not least in terms of obtaining sufficient sample size in surveys to make sound statistical inferences (Thompson, 2004). For example, one of IUCN's criteria for listing a population as Critically Endangered is evidence that the population numbers <250 mature individuals (Standards and Petitions Working Group, 2006). This is a well-known conservation challenge, because generally, the smaller the population, the harder it is to estimate its abundance from sampling methods (Thompson, 2004). When working in the marine environment, statistical problems may be exacerbated by logistical issues and financial constraints of ship time. This does not bode well for conservation and management of many marine species, particularly cetacean populations, whose size is small and whose conservation status is unfavourable (Perrin, 1999; Read and Wade, 2000; Sinha, 2002; Rojas-Bracho et al., 2006).

The difficulty of providing cost-effective estimates of absolute abundance has inspired researchers to develop a range of creative 'index methods' to quantify relative abundance instead (Aragones et al., 1997). Interviews, questionnaire surveys, land-based monitoring, reporting networks for opportunistic sightings, and beach surveys for stranded animals are attractive to researchers with limited financial resources, and trends can be inferred from long-term data on relative density, presence versus absence, or habitat use. No criticism is intended here. When trying to protect a population that is suspected to be small, it may be more important and cost-effective to implement conservation actions than to divert valuable resources into obtaining a robust abundance estimate (Jaramillo-Legoretta et al., 2007; Chades et al., 2008). Also, precise index methods may offer more statistical power to detect trends than imprecise abundance estimates (Thompson, 2004). However, index methods rely on the often untested and questionable assumption that the index is linearly related to absolute abundance, and that this relationship is constant over space or time. Ultimately, index methods are not meant to serve as replacements for estimates of

* Corresponding author. Address: Marine Mammal Research Unit, University of British Columbia, 2202 Main Mall, Room 247, AERL, Vancouver, BC, Canada V6T 1Z4. Tel.: +1 250 974 7103.
E-mail addresses: rmcw@st-and.ac.uk, pearseisland@hotmail.com (R. Williams).

0006-3207/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.biocon.2008.12.028

absolute abundance. Our concern here is two-fold: reliance on relative abundance indices may overstate the cost and difficulty of conducting systematic surveys, and underestimate the importance of estimating absolute abundance. Over the last few decades of endangered species research, it has become clear that abundance, as difficult as it can be to estimate, is actually one of the easiest demographic measures to provide (Taylor et al., 2000) as well as the one that is most essential for managing populations (Cooke, 1995). Indeed, abundance data are so important that some governments mandate their collection relative to anthropogenic mortality (Wade, 1998). Abundance estimates are also essential to modelling extinction risk, which is particularly high for small populations (Frankham, 2005; O'Grady et al., 2006).

While there is no universally ideal survey technique, the two most widely used techniques to estimate cetacean abundance are line transect surveys and related distance sampling methods (Buckland et al., 2001) and mark-recapture methods using photo-identification data (Hammond, 1986). These methods are used on a wide range of large, marine vertebrates, such as whale sharks (Meekan et al., 2006), polar bears (Marques et al., 2006), dugongs (Hines et al., 2005) and blue and humpback whales (Calambokidis and Barlow, 2004). Surveys to estimate abundance need not be expensive to be useful. Between the extremes of large-scale national, multinational or intergovernmental surveys (Branch and Butterworth, 2001; Hammond et al., 2002; Zerbini et al., 2007) and platform-of-opportunity surveys (Williams et al., 2006), there are cost-effective, small-boat surveys to estimate cetacean abundance using line transect (Dawson et al., 2004) and mark-recapture (Read et al., 2003) methods. Small-boat surveys are widely used, but they can carry problems: restriction to nearshore waters; small teams of observers; and a narrow field of view, which can reduce the number of sightings compared to higher platforms. The logistics of isolating observers in confined spaces make it difficult to set up double-platform line transect protocols on small boats to estimate trackline detection probability or to detect animals prior to responsive movement taking place. However, one perceived drawback of small boats is actually an advantage – the inability for small boats to work in rough seas means that search effort takes place under relatively good sighting conditions. Overall, if low-cost, small-boat surveys can produce unbiased abundance estimates, then this gives funders and decision-makers incentive to fund such surveys for the many regions of the world where such data are lacking and needed.

A broad distinction can be made between design-based and model-based methods for estimating abundance, where the former relies on properties of the survey design and the latter on a statistical model of the animal density (Buckland et al., 2004). Design-based methods require random placement of tracklines, but give unbiased estimates; model-based methods allow estimation from non-randomized surveys (such as platforms of opportunity) but can give unreliable results if survey effort is not spread through the study area, and several technical issues remain to be addressed with their use. The focus of this paper is on the efficacy of small-boat, design-based surveys in which the cost of the platform is low, but the statistical principles underlying design and analysis are conventional.

Recent reviews for small-boat sightings surveys are available on good design (Thomas et al., 2007) and field protocols (Dawson et al., 2008), and those topics are not duplicated here. In contrast, few studies have had the opportunity to evaluate the efficacy of small-boat surveys. The 'resident', fish-eating killer whale (*Orcinus orca*) populations of the northeast Pacific provide one real-world opportunity to evaluate the reliability of small-boat surveys. These whales have been studied for decades in British Columbia (BC) and Washington State (WA) (Bigg, 1982; Ford et al., 2000; Williams and Lusseau, 2006), and every killer whale in the northern resident population has been identified in most years during annual censuses (Cetacean Research Program, Pacific Biological Station, Fisheries and Oceans Canada). A systematic line transect survey was recently designed (Thomas et al., 2007) and conducted (Williams and Thomas, 2007) to estimate abundance of six cetacean species in BC coastal waters. Small boats were employed, costing about US$1000 per day. By contrast, ship time for large surveys may easily cost an order of magnitude higher than this, and traditionally represents the largest cost in a sightings survey. The northern resident killer whale population is used here as a case study in which line transect and mark-recapture abundance estimates are compared to the known population size.

2. Methods

2.1. Study design and field methods

A design-unbiased, stratified survey was planned using the automated survey design algorithm in Distance 4.1 (Thomas et al., 2004) for BC's coastal waters. The core area for the population, Queen Charlotte Basin including Queen Charlotte Strait, was surveyed twice: in summers 2004 and 2005. The other area, Johnstone Strait, was surveyed by line transect only in summer 2004, but was transited in 2005 and identification photographs taken. Standard line transect protocols were followed in both seasons (Williams and Thomas, 2007). When a sighting was made, the data recorder noted radial distance, radial angle (measured using angle boards), time, location, species and school size. Radial distances were either measured with reticles or photogrammetry; if a visual estimate had to be made, then estimates were corrected using observer-specific distance estimation experiments (Williams and Thomas, 2007; Williams et al., 2007).

For each killer whale encounter, the designed survey was suspended and we 'closed' on the whales to collect identification photographs (left side dorsal fin and saddle patch) and accurate estimates of school size. Images were sent to the Cetacean Research Program, Pacific Biological Station for comparison with their long-term catalogues (Ford et al., 2000; Ellis et al., 2007), and those photographs of good quality (Wilson et al., 1999) that gave certain matches due to unique markings were included in the analysis. This special case, in which all animals can be recognised, is unique (Ellis et al., 2007), but was essential for comparing our estimates to known population size (although 2005 was an incomplete census year). More commonly, one estimates the proportion of unmarked animals in the population and adjusts abundance estimates upwards (Wilson et al., 1999).

2.2. Analysis of line transect survey data

The analysis methods closely followed those described by Williams and Thomas (2007), with two differences. Firstly, instead of selecting one detection function model and basing inferences on that model, we used a weighted average of many plausible models (Buckland et al., 1997; Burnham and Anderson, 2002). Secondly, instead of using parametric methods to estimate variance, we used a nonparametric bootstrap, which makes fewer assumptions (Buckland et al., 2001). To increase sample sizes, killer whales seen during transit legs (i.e., off pre-determined tracklines but where observers surveyed as if they were on pre-determined tracklines) were included in all bootstrap replicates for detection function modeling and estimating mean school size, but were not used in calculating encounter rate. Also, schools identified as 'transient'-type killer whales (i.e., members of the sympatric, mammal-eating population, Ford et al., 2000) were used in detection function modeling, but not for estimating mean school size or encounter rate.

This assumed that the detection function for resident and transient types was the same.

Perpendicular distance data were right-truncated, and several standard detection function models (Buckland et al., 2001) were fitted to the data using Distance 5.0 (Thomas et al., 2006). Models with more than two parameters were not considered due to the small number of observations available. One model that produced a fit judged to be implausible for the species and survey was excluded from the candidate set (see Section 3). All others were included in the bootstrap analysis. Note that, as is good practice, density and abundance were not calculated at this stage. The analyst (LT) was unaware of true population size prior to conducting the analysis, so that knowledge of the consequences of different detection function fits could not influence model selection.

For the bootstrap analysis, 10,000 bootstrap resample datasets were generated by sampling with replacement from the transect lines within each stratum. The choice of 10,000 replicates was more than typically performed, but was required to obtain estimates and confidence intervals accurate to three significant figures, and took only a few hours of computer time. For each resample dataset, all candidate detection function models were fitted and the model with the lowest Akaike Information Criterion (AIC) value was selected. The selected model was used to estimate mean probability of detecting a school, given that it was within the truncation distance. Mean school size was also calculated within each resample data set, by fitting a least-squares regression of the logarithm of school size on estimated detection probability at the distance the school was sighted, and predicting school size at a detection probability of 1 (i.e., zero distance). Density and abundance were then estimated for each resampled dataset by first estimating density and abundance in each stratum, and then taking a weighted average (Williams and Thomas, 2007). While detection function model averaging can be conducted in program Distance, we were unable to do so in this case due to geographic and temporal stratification. Therefore, the bootstrap analysis was performed in R (R Development Core Team, 2007), calling Distance only to fit the detection functions for each bootstrap resample.

Estimates of detection probability, population mean school size, density and abundance were taken as the mean of the bootstrap resample estimates (Buckland et al., 1997). Coefficients of variation (CVs) were calculated as standard deviation of the bootstrap estimates divided by the mean, and the percentile method (Buckland et al., 2001) was used to obtain confidence intervals.

2.3. Mark-recapture analysis from photographic data

Photographs from 2004 were treated as the first sampling occasion, and those from 2005 were treated as the second. Consequently, analysis options were restricted to a two-sample estimator, which has been used extensively for studies of cetacean populations (Hammond, 1986; Read et al., 2003; Calambokidis and Barlow, 2004). The population was assumed to be closed between sampling events, which is reasonable for this long-lived species. Chapman's modified two-sample (Lincoln–Petersen) estimator for small sample sizes and its associated log-normal 95% confidence intervals were used (Seber, 1982; Hammond, 1986).

3. Results

3.1. Line transect estimate

One thousand and four hundred and sixteen nautical miles (2715 km) of trackline was surveyed in 2004 (from 6 to 21 June, and from 4 July to 15 August) and 917 nm (1698 km) in 2005 (from 1 to 29 August). Eighteen killer whale schools were recorded during the survey; of these, 15 were resident and three were transient. Four of these sightings were made during transit-leg surveys. A truncation distance of 1500 m was chosen, in that it only required dropping two observations, but provided a good fit to the data for most detection function models. Most models produced estimates of average detection probability in the range 0.5–0.75 (Table 1): two exceptions were the uniform key function with no adjustments, which assumes detection probability to be 1, and the hazard rate key function, which fitted a steeply declining detection function that estimated detection probability to be 0.5 at around 100 m, and average detection probability to be only 0.17. Although the former (certain detection of schools out to 1500 m) is conceivable given the species, group sizes and good sighting conditions during the surveys, the latter (average detection probability 0.17) is implausible for this species and was excluded from consideration in the bootstrap analysis. The mean detection probability, averaging over all bootstrap replicates, was 0.58 (Fig. 1; CV = 33.7%; 95% CI = 0.30–1.00).

Table 1
Fitted detection function models, estimated average detection probabilities ($\hat{P}_a$) and CVs ($CV(\hat{P}_a)$), Kolmogorov–Smirnov goodness of fit *p*-values (KS-*p*), Akaike Information Criterion values (AIC), proportion of the bootstrap resamples for which this model was selected (w_{boot}) and mean of the estimated detection function probabilities for the bootstraps where this model was selected ($\hat{P}_{a,boot}$).

Model[a]	$\hat{P}_a$	$CV(\hat{P}_a)$	KS-*p*	AIC	w_{boot}	$\hat{P}_{a,boot}$
A. unif	1.00	0.00	0.02	234.02	0.14	1.00
B. unif+cos(1)	0.58	0.13	0.28	230.41	0.32	0.60
C. unif+cos(1)+cos(2)	0.49	0.23	0.41	231.86	0.06	0.45
D. unif+poly(2)	0.73	0.10	0.09	232.70	0.28	0.52
E. unif+poly(2)+poly(4)	0.56	0.24	0.32	232.02	0.01	0.48
F. hn	0.56	0.17	0.31	230.91	0.07	0.38
G. hn+cos(2)	0.46	0.26	0.45	232.06	0.13	0.36
H. hn+Herm(4)	0.56	0.27	0.31	232.89	0.00	–[b]
I. hr	0.17	0.71	0.56	228.91	–[c]	–
J. Model averaged	–	–	–	–	–	0.58

[a] Key functions: unif = uniform; hn = half normal; hr = hazard rate. Series expansions: cos(x) = cosine of order x; poly(x) = simple polynomial of order (x); Herm(x) = Hermite of order x. Formulae for these functions are given in Buckland et al. (2001), p.47.

[b] This model was not selected during the bootstrap.

[c] This model was excluded from the candidate set for bootstrapping; see text for details.

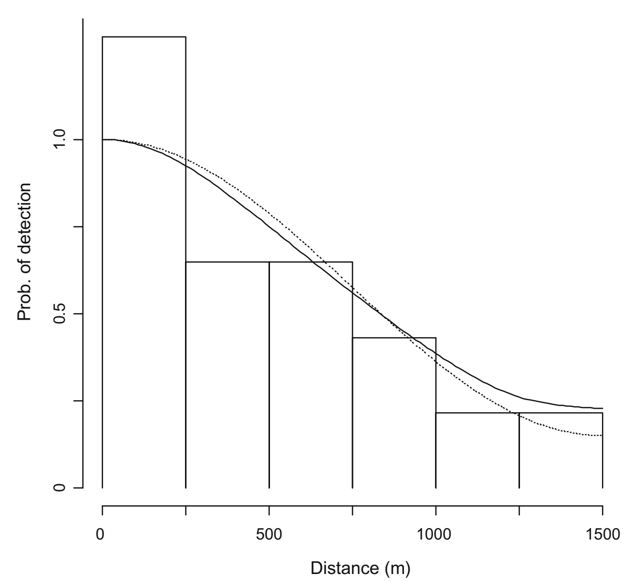

Fig. 1. Histogram of observed distances from line transect survey, with model-averaged detection function (solid line) and detection function from the half-normal model used by Williams and Thomas (2007, dotted line) superimposed.

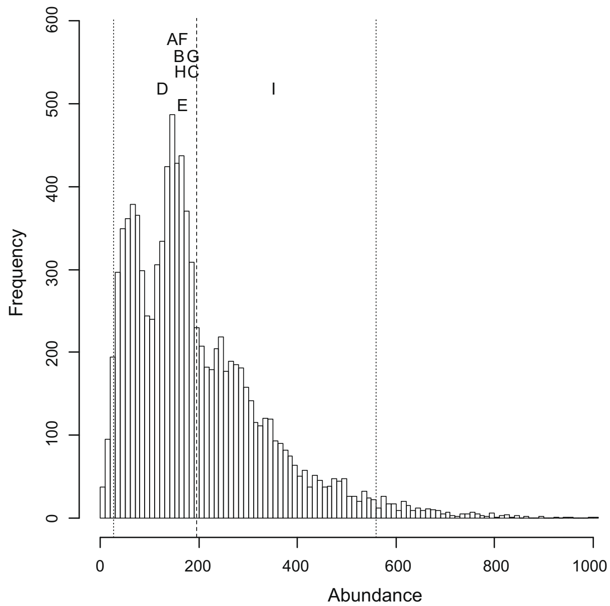

Fig. 2. Histogram of 10,000 bootstrap estimates of abundance from the analysis of line transect data. The dashed vertical line shows the mean of the bootstrap estimates (i.e., the model averaged estimate of abundance) and the dotted lines show the 2.5th and 97.5th percentile of the bootstrap estimates. The horizontal positioning of the letters A-I indicates the estimates that would have been obtained by assuming each of the detection function models listed in Table 1 was the correct model (see Table 1 for functional form of each model).

The mean resident group size observed in the survey was 5.07 (CV = 28.2%), with a minimum, median and maximum of 1, 3, and 25, respectively. The regression-based estimate of population mean group size, averaged over bootstrap resamples was 2.69 (CV = 36.1%; 95% CI = 1.64–5.58).

Mean density overall, averaged over bootstrap resamples, was 0.00808 animals/nm^2 (CV = 71.0%; 95% CI = 0.00111–0.0230). Density in Queen Charlotte Basin was similar in 2004 (0.00621; CV = 78.9%; 95% CI = 0–0.0182) and 2005 (0.0112; CV = 140%; 95% CI = 0–0.0451), but was highest in Johnstone Strait (0.310; CV = 55.4%; 95% CI: 0.0597–0.722). These densities correspond to an average of 160 animals in Queen Charlotte Basin (CV = 82.3%; 95% CI = 0–501) and 38 in Johnstone Strait (95% CI = 7–88) for a total abundance estimate of 197 animals (Fig. 2; 95% CI = 27–559).

3.2. Mark-recapture estimate

Fourty-three individuals were identified from photographs in 2004, and 59 in 2005, with 10 identified in both years. This corresponds to a Chapman abundance estimate of 239 (i.e., $(44 * 60/11)-1$) with log-normal, 95% CI of 154–370 and CV = 22.7%.

4. Discussion

Our cost-effective case study to estimate abundance of a naturally small population demonstrates that both methods worked well under real-world conditions. Model-averaged line transect (195, 95% CI: 30–564) and mark-recapture (239, 95% CI 133–345) methods performed well at estimating the true size of the northern resident killer whale population, which was 219 animals in 2004 and 235 in 2006 (the year after our study; both estimates courtesy Cetacean Research Program, DFO). The mark-recapture and line transect estimates are not directly comparable, because the former estimates population size in the study area on the days that the survey took place, while the latter estimates the (larger) size of the super-population using the study area during the survey interval (Clambokidis and Barlow, 2004). Nevertheless, because we covered the core area for the population of northern resident killer whales at the time of year of the survey (Ford et al., 2000), the two estimates should be similar. Line transect and mark-recapture methods are complementary, so studies that use one approach can provide ancillary information that benefits studies that use the other (Cassey and Ussher, 1999; Calambokidis and Barlow, 2004; Evans and Hammond, 2004). Useful results were obtained in a relatively short time frame (42 days of ship time in 2004, and 28 days in 2005). This is important, because some cetacean populations may not have many years left for us to conduct long-term studies to assess population size, status and trends (Rojas-Bracho et al., 2006; Clapham and Van Waerebeek, 2007; Taylor et al., 2007; Taylor and Gerrodette, 1993).

Approximately 10% of the study budget was spent on survey design, which brought together a statistician, a geographer and a field biologist (Thomas et al., 2007). This approach facilitated relatively simple, conventional distance sampling analyses, and helped to satisfy the assumption of equal capture probability underpinning a simple two-sample mark-recapture estimator. Automated survey design algorithms are built into freely available software (Thomas et al., 2004, 2007), and should be used where possible. Even when one plans to use model-based distance sampling methods, good coverage of the study area helps to satisfy assumptions about sampling the range of explanatory covariates that underpin spatial modelling methods (Hedley et al., 1999; Cañadas and Hammond, 2006; Williams et al., 2006). Clear and creative thinking about survey design is particularly important for difficult study areas such as linear habitats (e.g., rivers) and highly non-convex regions (such as fjords); some suggestions are given by Thomas et al. (2007).

For the line transect estimator, the detection function was modelled using fewer sightings than the recommended 60–80 (Buckland et al., 2001), but provided a reasonable fit to the data and concurred with our understanding of killer whale detectability. Importantly, the analyst conducted the detection function estimation as a worthy analysis in its own right (rather than a step along the way to abundance estimation), only moving on to density and abundance estimation once the final model or set was chosen. A relatively automatic approach of fitting several candidate models and selecting among them with AIC worked well, but unthinking reliance on model selection statistics is not recommended generally. Common sense and a good understanding of the biology of the species are also needed, so that analysis proceeds in the form of a dialogue between analyst and biologist. It was reasonable to assume that all animals on the trackline were seen for this conspicuous species, a fundamental assumption of conventional line transect methods, but this may be unreasonable for cryptic species or under poorer sighting conditions. Model-averaging to incorporate detection function uncertainty is a good way to cope with the small sample sizes inherent in studies of rare or elusive animals (Thompson, 2004). The model-averaged estimate of mean detection probability (0.58) was similar to that from the best-fitting model (0.56); but the CV was considerably larger (36% vs. 17%), reflecting the additional uncertainty due to model selection. Additional uncertainty did not affect the final CV of the density estimate dramatically (71.2% vs. 67.4%), because much of the uncertainty in density comes from encounter rate variation. A bootstrap approach to address model uncertainty will be particularly useful for studies of endangered species when paucity of data precludes testing for failure of model assumptions, but care needs to be taken to ensure numerical stability by constraining to fit only a few parameters at most and restricting the candidate model set to those that produce plausible results.

Mark-recapture estimates from surveys such as this would be expected to show negative bias due to heterogeneity in capture probability (Hammond, 1986; Wilson et al., 1999). Our study was also at the limit of the minimum 7–10 recaptures suggested for mark-recapture analysis (Seber, 1982). Nevertheless, the point estimate was very close to true population size. The apparent precision of the mark-recapture estimate was artificially improved by violation of underlying assumptions, including the assumption of demographic closure between 2004 and 2005 (Wilson et al., 1999). This is important in conservation studies, because a negatively biased variance estimate can overestimate sustainable limits to anthropogenic mortality (Wade, 1998). Of course, there are few options available to researchers who only have funding to conduct one or two surveys. Clearly, two surveys are the minimum needed to conduct mark-recapture experiments, but one generally should prioritise good survey design and field protocols over increasing number of sampling occasions, and be sure to allocate adequate resources to processing and analysis of potentially sparse data (Read et al., 2003). The overarching goal is to increase capture probability and number of recaptures overall (Seber, 1982; Hammond, 1986; Wilson et al., 1999). As funding increases, and resulting number of sampling occasions and resightings increase, researchers can employ more sophisticated models to obtain better information about the population under study (e.g., Pollock, 1982; Durban and Elston, 2005).

To some extent, studies to estimate abundance of rare species involve both scientific realities and managing expectations. Our intent is not to give anyone false hope that studies with small sample sizes can always be salvaged. Two things can go wrong: firstly, small sample sizes of observations mean that possible problems with the data cannot be diagnosed and simple models must be fit, potentially resulting in biased estimates; secondly, estimates will often be imprecise. The first problem can be overcome to some extent by careful survey design, field methods, and analysis. The extent to which this will work depends on the study species and environment. For the second, any desired level of precision can be gained with sufficient survey effort aided by appropriate stratification, and the required effort can be quantified based on pilot survey data (see Buckland et al., 2001 for line transects and Devineau et al., 2006 for mark-recapture). Nevertheless, financial and logistical constraints dictate that imprecise estimates are common features of studies on small populations. Our main message is that in many parts of the world, and for many understudied taxa, abundance estimates on the right order of magnitude, even with high variance, would represent a useful starting point for informing conservation strategies (Perrin, 1999). This is particularly the case for rare species, or populations that may be critically endangered, where the conservation stakes are especially high.

For many under-studied populations, a two-sample mark-recapture estimator that provides point estimates on the correct order of magnitude is a good starting point (Hammond, 1986), as would an abundance estimate from a line transect survey that yielded few sightings, as long as associated uncertainty is reported explicitly. In our experience, surveys to provide these data often are not conducted because of financial restrictions or pessimism. Here it is shown that reliable abundance estimates for a small population can be generated by conducting systematic surveys from a relatively inexpensive survey platform. Initial estimates of abundance from pilot studies such as these can be thought of not so much as definitive as they are hypotheses to be tested as new data are collected. For example, a long-term study of individually recognisable dolphins in Moray Firth (Scotland) began with a land-based census to provide a minimum count (Hammond and Thompson, 2001). If built upon a solid base, population assessment programs can improve and evolve through time, producing increasingly reliable abundance estimates as sample sizes allow more sophisticated analytical methods, and improved survey methods are developed. Tentative estimates can be reported along the way with discussions about their limitations and associated measures of statistical confidence.

Acknowledgements

We thank Raincoast Conservation Society, its funders and field observers (especially O. Andrews, B. Falconer and S. Heinrich) for assistance to R.W. with data collection. Graeme Ellis, Cetacean Research Program, Pacific Biological Station, Fisheries and Oceans Canada, Nanaimo BC Canada identified individual killer whales from photographs, and he and John Ford encouraged the study initially and provided unpublished data from annual censuses for comparison. Humane Society International provided support for analyses. E. Ashe, S. Dawson, J. Durban, P.S. Hammond, J. Klaich, D. Lusseau and the anonymous reviewers provided valuable feedback. All whales were approached for photo-identification under SARA permit MML-12 to R.W.

References

Aragones, L.V., Jefferson, T.A., Marsh, H., 1997. Marine mammal survey techniques applicable in developing countries. Asian Marine Biology 14, 15–39.
Bigg, M., 1982. An assessment of killer whale (*Orcinus orca*) stocks off Vancouver Island, British Columbia. Report of the International Whaling Commission, vol. 32, pp. 655–666.
Branch, T.A., Butterworth, D.S., 2001. Estimates of abundance south of 60 S for cetacean species sighted frequently on the 1978/79 to 1997/98 IWC/IDCR-SOWER sighting surveys. Journal of Cetacean Research and Management 3 (3), 251–270.
Buckland, S.T., Burnham, K.P., Augustin, N.H., 1997. Model selection: an integral part of inference. Biometrica 53, 603–618.
Buckland, S., Anderson, D.R., Burnham, K.P., Laake, J., Borchers, D., Thomas, L., 2001. Introduction to Distance Sampling: Estimating Abundance of Biological Populations. Oxford University Press, Oxford.
Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., Borchers, D.L., Thomas, L., 2004. Advanced Distance Sampling. Oxford University Press, Oxford.
Burnham, K.P., Anderson, D.R., 2002. Model Selection and Inference: A Practical Information Theoretic Approach, second ed. Wiley.
Calambokidis, J., Barlow, J., 2004. Abundance of blue and humpback whales in the Eastern North Pacific estimated by capture–recapture and line-transect methods. Marine Mammal Science 20 (1), 63–85.
Cañadas, A., Hammond, P.S., 2006. Model-based abundance estimates for bottlenose dolphins (*Tursiops truncatus*) off southern Spain: implications for management. Journal of Cetacean Research and Management 8 (1), 13–27.
Cassey, P., Ussher, G.T., 1999. Estimating abundance of tuatara. Biological Conservation 88 (3), 361–366.
Chades, I., McDonald-Madden, E., McCarthy, M.A., Wintle, B., Linkie, M., Possingham, H., 2008. When to stop managing or surveying cryptic threatened species? Proceedings of the National Academy of Sciences 105, 13936–13940.
Clapham, P., Van Waerebeek, K., 2007. Bushmeat and bycatch: the sum of the parts. Molecular Ecology 16, 2607–2609.
Cooke, J.G., 1995. The international whaling commission's revised management procedure as an example of a new approach to fishery management. In: Blix, A.S., Walløe, L., Ulltang, Ø. (Eds.), Whales, Seals, Fish and Man. Elsevier, Amsterdam, pp. 647–657.
Dawson, S.M., Slooten, E., DuFresne, S.D., Wade, P.R., Clement, D.M., 2004. Small-boat surveys for coastal dolphins: line-transect surveys of Hector's dolphins (*Cephalorhynchus hectori*). Fishery Bulletin 102 (3), 441–451.
Dawson, S., Wade, P., Slooten, E., Barlow, J., 2008. Design and field methods for sighting surveys of cetaceans in coastal and riverine habitats. Mammal Review 38 (1), 19–49.
Devineau, O., Choquet, R., Lebreton, J.D., 2006. Planning capture-recapture studies: straightforward precision, bias, and power calculations. Wildlife Society Bulletin 34 (4), 1028–1035.
Durban, J.W., Elston, D.A., 2005. Mark-recapture with occasion and individual effects: abundance estimation through Bayesian model selection in a fixed dimensional parameter space. Journal of Agricultural, Biological and Environmental Statistics 10 (3), 291–305.
Ellis, G.M., Ford, J.K.B., Towers, J.R., 2007. Northern Resident Killer Whales in British Columbia: Photo-Identification Catalogue 2007. Pacific Biological Station, Nanaimo.
Evans, P.G.H., Hammond, P.S., 2004. Monitoring cetaceans in European waters. Mammal Review 34, 131–156.
Ford, J.K.B., Ellis, G.M., Balcomb, K.C., 2000. Killer Whales, second ed. UBC Press, Vancouver.
Frankham, R., 2005. Genetics and extinction. Biological Conservation 126, 131–140.
Hammond, P.S., 1986. Estimating the size of naturally marked whale populations using capture–recapture techniques. Report of the International Whaling Commission (special Issue 8), pp. 253–282.

Hammond, P.S., Thompson, P.M., 2001. Minimum estimate of the number of bottlenose dolphins (*Tursiops truncatus*) in the Moray Firth. Biological Conservation 56, 79–87.

Hammond, P.S., Berggren, P., Benke, H., Borchers, D.L., Collet, A., Heide-Jørgensen, M.P., Heimlich, S., Hiby, A.R., Leopold, M.F., Øien, N., 2002. Abundance of harbour porpoise and other cetaceans in the North Sea and adjacent waters. Journal of Applied Ecology 39, 361–376.

Hedley, S.L., Buckland, S.T., Borchers, D.L., 1999. Spatial modelling from line transect data. Journal of Cetacean Research and Management 1 (3), 255–264.

Hines, E.M., Adulyanukosol, K., Duffus, D.A., 2005. Dugong (*Dugong dugon*) abundance along the Andaman coast of Thailand. Marine Mammal Science 21 (3), 536–549.

Jaramillo-Legoretta, A.M., Rojas-Bracho, L., Brownell Jr., R.L., Read, A.J., Reeves, R.R., Ralls, K., Taylor, B.L., 2007. Saving the vaquita: immediate action not more data. Conservation Biology 21, 1653–1655.

Marques, T.A., Andersen, M., Christensen-Dalsgaard, S., Belikov, S., Boltunov, A., Wiig, O., Buckland, S.T., Aars, J., 2006. The use of global positioning systems to record distances in a helicopter line-transect survey. Wildlife Society Bulletin 34, 759–763.

Meekan, M.G., Bradshaw, C.J.A., Press, M., McLean, C., Richards, A., Quasnichka, S., Taylor, T.J., 2006. Population size and structure of whale sharks *Rhincodon typus* at Ningaloo Reef, Western Australia. Marine Ecology Progress Series 319, 275–285.

O'Grady, J.J., Brook, B.W., Reed, D.H., Ballou, J.D., Tonkyn, D.W., Frankham, R., 2006. Realistic levels of inbreeding depression strongly affect extinction risk in wild populations. Biological Conservation 133, 42–51.

Perrin, W.F., 1999. Selected examples of small cetaceans at risk. In: Twiss, J.R., Reeves, R.R. (Eds.), Conservation and Management of Marine Mammals. Smithsonian Institution, Washington, pp. 296–310.

Pollock, K.H., 1982. A capture–recapture design robust to unequal probability of capture. Journal of Wildlife Management 46 (3), 752–757.

R Development Core Team, 2007. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>.

Read, A.J., Wade, P.R., 2000. Status of marine mammals in the United States. Conservation Biology 14 (4), 929–940.

Read, A.J., Urian, K.W., Wilson, B., Waples, D.M., 2003. Abundance of bottlenose dolphins in the bays, sounds and estuaries of North Carolina. Marine Mammal Science 19 (1), 59–73.

Rojas-Bracho, L., Reeves, R.R., Jaramillo-Legorreta, A., 2006. Conservation of the vaquita *Phocoena sinus*. Mammal Review 36 (3), 179–216.

Seber, G.A.F., 1982. The Estimation of Animal Abundance and Related Parameters, second ed. The Blackburn Press, New Jersey.

Sinha, R.K., 2002. An alternative to dolphin oil as a fish attractant in the Ganges River system: conservation of the Ganges River Dolphin. Biological Conservation 107, 253–257.

Standards and Petitions Working Group, 2006. Guidelines for Using the IUCN Red List Categories and Criteria. Version 6.2. Prepared by the Standards and Petitions Working Group of the IUCN SSC Biodiversity Assessments Sub-Committee in December 2006. <http://app.iucn.org/webfiles/doc/SSC/RedList/RedListGuidelines.pdf>.

Taylor, B.L., Gerrodette, T., 1993. The uses of statistical power in conservation biology: the vaquita and northern spotted owl. Conservation biology 7 (3), 489–500.

Taylor, B.L., Wade, P.R., DeMaster, D.P., Barlow, J., 2000. Incorporating uncertainty into management models for marine mammals. Conservation Biology 14 (5), 1243–1252.

Taylor, B.L., Martinez, M., Gerrodette, T., Barlow, J., Hrovat, Y.N., 2007. Lessons from monitoring trends in abundance of marine mammals. Marine Mammal Science 23 (1), 157–175.

Thomas, L., Laake, J.L., Strindberg, S., Marques, F.F.C., Buckland, S.T., Borchers, D.L., Anderson, D.R., Burnham, K.P., Hedley, S.L., Pollard, J.H., Bishop, J.R.B., 2004. Distance 4.1. Release 2. Research Unit for Wildlife Population Assessment, University of St. Andrews, UK. <http://www.ruwpa.st-and.ac.uk/distance/>.

Thomas, L., Laake, J.L., Strindberg, S., Marques, F.F.C., Buckland, S.T., Borchers, D.L., Anderson, D.R., Burnham, K.P., Hedley, S.L., Pollard, J.H., Bishop, J.R.B., Marques, T.A., 2006. Distance 5.0. Release 2. Research Unit for Wildlife Population Assessment, University of St. Andrews, UK. <http://www.ruwpa.st-and.ac.uk/distance/>.

Thomas, L., Williams, R., Sandilands, D., 2007. Survey design for complex survey regions. Journal of Cetacean Research and Management 9 (1), 1–13.

Thompson, W.L. (Ed.), 2004. Sampling Rare or Elusive Species: Concepts, Designs and Techniques for Estimating Population Parameters. Island Press, p. 429.

Wade, P.R., 1998. Calculating limits to the allowable human-caused mortality of cetaceans and pinnipeds. Marine Mammal Science 14 (1), 1–37.

Williams, R., Lusseau, D., 2006. A killer whale social network is vulnerable to targeted removals. Biology Letters 2 (4), 497–500.

Williams, R., Thomas, L., 2007. Distribution and abundance of marine mammals in coastal waters of British Columbia, Canada. Journal of Cetacean Research and Management 9 (1), 15–28.

Williams, R., Hedley, S.L., Hammond, P.S., 2006. Modeling distribution and abundance of Antarctic baleen whales using ships of opportunity. Ecology and Society 11 (1), 1. <http://www.ecologyandsociety.org/vol11/iss1/art1/>.

Williams, R., Leaper, R., Zerbini, A., Hammond, P.S., 2007. Methods for investigating measurement error in cetacean line transect surveys. Journal of the Marine Biological Association of the United Kingdom 87, 313–320.

Wilson, B., Hammond, P.S., Thompson, P.M., 1999. Estimating size and assessing status of a coastal bottlenose dolphin population. Ecological Applications 9, 288–300.

Zerbini, A.N., Waite, J.M., Durban, J.W., LeDuc, R., Dahlheim, M., Wade, P., 2007. Estimating abundance of killer whales in the nearshore waters of the Gulf of Alaska and Aleutian Islands using line-transect sampling. Marine Biology 150, 1033–1045.