

# Monitoring population dynamics of right whales off southern Australia

Project 2009/41: Final Report  
to The Australian Marine Mammal Centre

13<sup>th</sup> April, 2011.

J.L. Bannister<sup>\*</sup>, S.L. Hedley, M.V. Bravington and S.R. Burnell<sup>\*</sup>

<sup>1</sup> Western Australian Museum, Locked Bag 49, Welshpool DC, WA 6986.

<sup>2</sup> The School House, Denhead, St Andrews, Fife, KY16 8PA, United Kingdom.

<sup>3</sup> CSIRO Mathematical & Information Sciences, Marine Lab., Castray Esplanade, Hobart, TAS 7001.

<sup>4</sup> South Australian Museum, North Terrace, Adelaide, SA 5000.

---

<sup>\*</sup> NB. While Bannister had overall management of the project, and he and Burnell provided the data on which it was based, the analysis and findings are the joint responsibility of Hedley and Bravington.

## ABSTRACT

Southern right whales have been systematically surveyed in their breeding grounds off southern and western Australia since the early 1990s. The survey at the Head of Bight has been ongoing every year since 1991. Counts and photo-identification records are collected from a shore-based lookout. Aerial surveys along the coastline from Cape Leeuwin, WA, to Ceduna, SA, have been conducted annually since 1993, yielding 'census' counts of the population present in the area each year, together with photographs suitable for identification of individuals. There is a strong cohort-structure induced on the population by the whales' typical three-year breeding cycle, and this, together with occasional preferences to skip-breed (giving a four-year breeding interval), leads to considerable inter-annual fluctuations in the counts, despite a rather steady trend over the longer term. In principle the photo-ID data could help explain the fluctuations in the counts, and thus remove some of the imprecision in the count estimates.

Current count data from the aerial survey indicate that the population is increasing at a healthy rate, *ca* 8% p.a., with a somewhat lower rate (*ca* 5.4% p.a.) for the Head of Bight site. This project aimed to address the question of whether two sets of annual surveys were needed, by examining how the different sources of data might be used to complement each other, and investigating the potential loss in information and precision under monitoring strategies with reduced frequency. In particular, the work undertaken aimed to develop a long-term monitoring strategy that would in future allow both efficient monitoring of population trend and abundance, and studies of linkages between population dynamics and environmental changes in the whales' feeding grounds (e.g. Leaper *et al.*, 2006).

We developed a multistate Mark-Recapture (MSMR) model that could be applied to the photo-ID data from either survey, and that could be combined with the count data. Rather substantial differences in the population dynamics were noted between the Head of Bight site, and the WA and SA survey sites (WASA) as a whole, with longer calving intervals at the Head of Bight. Given the differences in population dynamics, it was apparent that we couldn't easily use Head of Bight photo-ID data to help explain the fluctuations in the WASA counts, and therefore we concentrated on WASA photo-ID and count data alone. We wrote bespoke software for doing this, estimating annual probabilities of (a) successfully obtaining a photo-ID quality image of a cow-calf pair, and (b) skip-breeding, as random-effects, and other parameters (such as death rate) as fixed effects.

The parameter estimates from the models were used to simulate future data, allowing different trends in recruitment so that the simulated future population trajectories could diverge from what has been seen to date. A variety of monitoring strategies were tested on each simulated dataset, in terms of the precision of abundance estimates. The results demonstrate how, in surveys with gap-years, there is a delay in recovering levels of precision in subsequent years of consecutive surveys, which may have implications for management of the population. We examined in detail two particular population trajectories: one for which the population continued to increase at a rate of about 8%, and one for which recruitment to the population averaged only about 5%. Under these circumstances, our results suggested that a minimum of four years of consecutive surveys should be undertaken in order to be reasonably confident of detecting a decline in the rate of increase of this magnitude. This result should not, of course, be generalised to other trajectories; in order to detect a more moderate decline in recruitment, increasingly higher numbers of consecutive survey years are required. The software that has been developed for this project is available to investigate such scenarios.

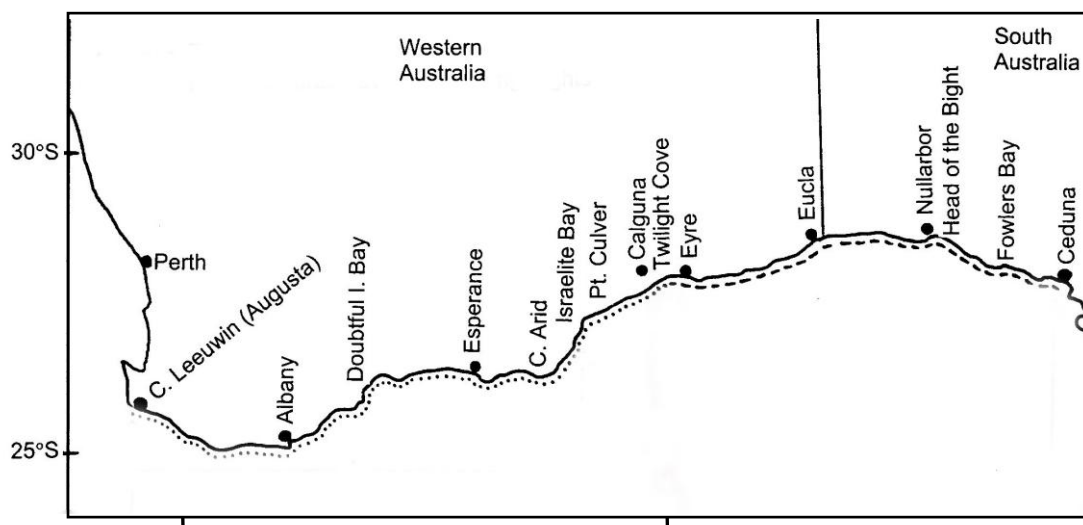
Coupled with strategies for abundance monitoring, we note that in order to detect possible linkages between population dynamics and environmental events, for example in the Southern Ocean, the ability of the model to detect 'skip-breeding' years, is key. Both Head of Bight and

WASA data are useful in this respect. Although not reported in detail, estimated precision in skip-probability is substantially lower for gappy survey strategies; annual surveys are recommended in this regard.

## 1: BACKGROUND

The ‘Australian’ right whale population comprises two distinct elements, each found close to the coast in winter/spring. The ‘western’ population occurs on the southern coast largely between Cape Leeuwin, WA and Ceduna, SA; the ‘eastern’ population occurs on the east coast. The former occurs in relatively large numbers, with an estimated current population size of ca 2900 individuals (Bannister, 2011); the latter in very small numbers, probably ca 500.

Aerial surveys of the ‘western’ element have been undertaken annually since 1976. These surveys provide information on both the number of whales present (‘counts’) and on individual histories (via photo-identification). The first series of surveys were carried out along the southern WA coast, but since 1993, they were extended into SA waters as far east as Ceduna (Figure 1). The series from 1993 was designed to provide statistically significant information on population size and trend over a fifteen year period (to include five three-year breeding cycles), i.e. to 2007 inclusive. An anomalously low count, particularly of breeding females, in 2007 (Bannister 2008a) led to subsequent surveys in 2008-2010 (Bannister 2008b; 2011).



**Figure 1:** Southern right whale surveys, southern Australia, 1976-2010. Dotted line indicates survey flights from 1976-1992. Dashed line indicates eastern extension to the survey from 1993 onwards.

Along both the WA and SA coastlines, extensive whaling occurred during the 19<sup>th</sup> century (Bannister, 2001). At one location, the Head of Bight, SA, whaling was not undertaken, and the ‘sub-population’ there is generally considered essentially pristine. Land-based counts and photo-identification studies have

been undertaken at the Head of Bight since 1990, with consistent protocols since 1991.

## 2: AIMS OF THIS PROJECT

This project aimed to address two objectives that might be asked of a monitoring program for southern right whales: (i) to monitor the population size, and (ii) to use fluctuations in the observation patterns to infer inter-annual variations in phenomena in the Southern Ocean (see Leaper *et al.* 2006; Pirzl *et al.*, 2008).

The fundamental tool for monitoring population size is the count data from the aerial survey off Western and South Australia (hereafter called the 'WASA' survey).

We assume for the moment (but also see section 3.3) that the count of cow-calf pairs is an unbiased index of the number of females breeding that year. There is an almost salmon-like 'cohort structure' induced by the typical 3-year breeding cycle. In some years, however, many females that are 'due' to breed in a particular year actually avoid breeding ('skip-breed') until the subsequent year, leading to unpredictable fluctuations in the cohort structure. Interpreting what the consequential (and rather large) inter-annual fluctuations in the count data mean for variation in underlying abundance is difficult. Use of the count data *alone* as a monitoring index for southern right whales is for example, much less informative than for, say, eastern Australian humpback whales. Even with a continuous annual survey series, the year-to-year variability is high because of skip-breeding and the 'cohort effect'. There is a negative correlation in the counts (if females fail to be counted in one year when they are expected to breed, they probably will the next), and hence a simple regression estimate overestimates uncertainty.

Furthermore, a reduced schedule for aerial surveys (i.e. no longer annual) would present some complications in interpreting the count in the year following a gap in the series. For example, a 'typical' count could be explained equally well either by a typical skip-breeding proportion in the gap year followed by a typical recruitment of new breeders, or by a high proportion of skip-breeding in the gap year and then an unusually low recruitment. While these two explanations would result in similar count data, they would have different implications for abundance.

Using the count data alone, there is little that can be done quantitatively to improve on the simple regression estimates. However, the photo-ID data give us a handle both on the cohort structure and on the extent of skip-breeding in different years. Intuitively, this might be useful for smoothing out the noise in

the count data to give a more stable abundance index; in the example of the previous paragraph, if we discovered that the skip-breeding rate in the gap year was typical, then we could rule out the ‘high-skip + low-recruitment’ interpretation of the count data.

Photo-ID data are available both from the WASA survey (though not all cow-calf pairs are photographed to database quality) and from the Head of Bight land-based survey (hereafter called the HoB survey<sup>1</sup>). If the population dynamics in the HoB data - and in particular the year-to-year skip-breeding - were similar enough to those in WASA as a whole, then we had thought that it might have been possible to use photo-ID data collected every year at the HoB to help smooth the count data from WASA, even when there are gaps in the WASA photo-ID and count data. In the event, our results concurred with previous analyses (Burnell, 2008; Bannister, 2009) suggesting appreciably different population dynamics at the HoB compared to WASA as a whole (see section 3.4).

The models we describe below can incorporate not only breeding females, but also other portions of the population such as unaccompanied adults and males. Our exploratory analyses suggested that only about 8% of non-breeders are present and get photographed in any one year; thus, the quite considerable extra model complexity was not justified for the main purpose of this work - assessing precision of estimates under different survey strategies. Therefore, in this report, we have concentrated on the abundance of adult females<sup>2</sup> (whether or not they have a calf in the current year), since this is the best-estimated sector of the population and should give the clearest indication of trend.

The remainder of the report contains:

- a description of the model(s) fitted to the WASA (count and) photo-id data;
- justification for not fitting a joint model to the HoB and WASA data;
- simulation results;
- a brief description the customized software developed for this project (noting it may be used to examine alternative scenarios in the future);
- conclusions from the simulations, and some recommendations.

We begin by describing the MSMR model fitted to photo-id data (not yet incorporating the count data).

---

<sup>1</sup> Count data are also available from the HoB, but these are not a focus of this project, since they cover only a subset of the population.

<sup>2</sup> i.e. females with a calf in the current year, or in a previous year.

### 3: A SINGLE-SITE MODEL FOR THE PHOTO-ID DATA

Before describing the model, it is perhaps worthwhile to briefly describe what the photo-id data actually look like. Typical example photo-id records are given below, together with likely - but not definitive - interpretations:

		Year																						
Individual ID		B	.	.	B	.	.	.	B	.	.	B	.	.	B	.	.	B	.	.	B	'textbook', mostly 3-year gaps with one 4-year gap		
																						B	3-year gap, 2-year gap, 4-year gap; breeding is not always regular	
																							B	3-year gap; 3-year gap but present in the WASA survey area in a non-calving year; 3-year gap then recorded as UA not CC; 2-year gap.
																								B

For the WASA photo-ID model, we have not distinguished between different sites within WASA, as intra-WASA movements do not affect trend or skip-breeding estimates (our analyses in section 3.3 confirm that the breeding female count does measure up as a true census). Exactly the same structure can be applied to HoB data too, albeit with a slightly different interpretation. We did not, however, attempt to deal with whales seen sometimes in HoB and sometimes elsewhere, since the datasets have not been fully formally cross-matched.

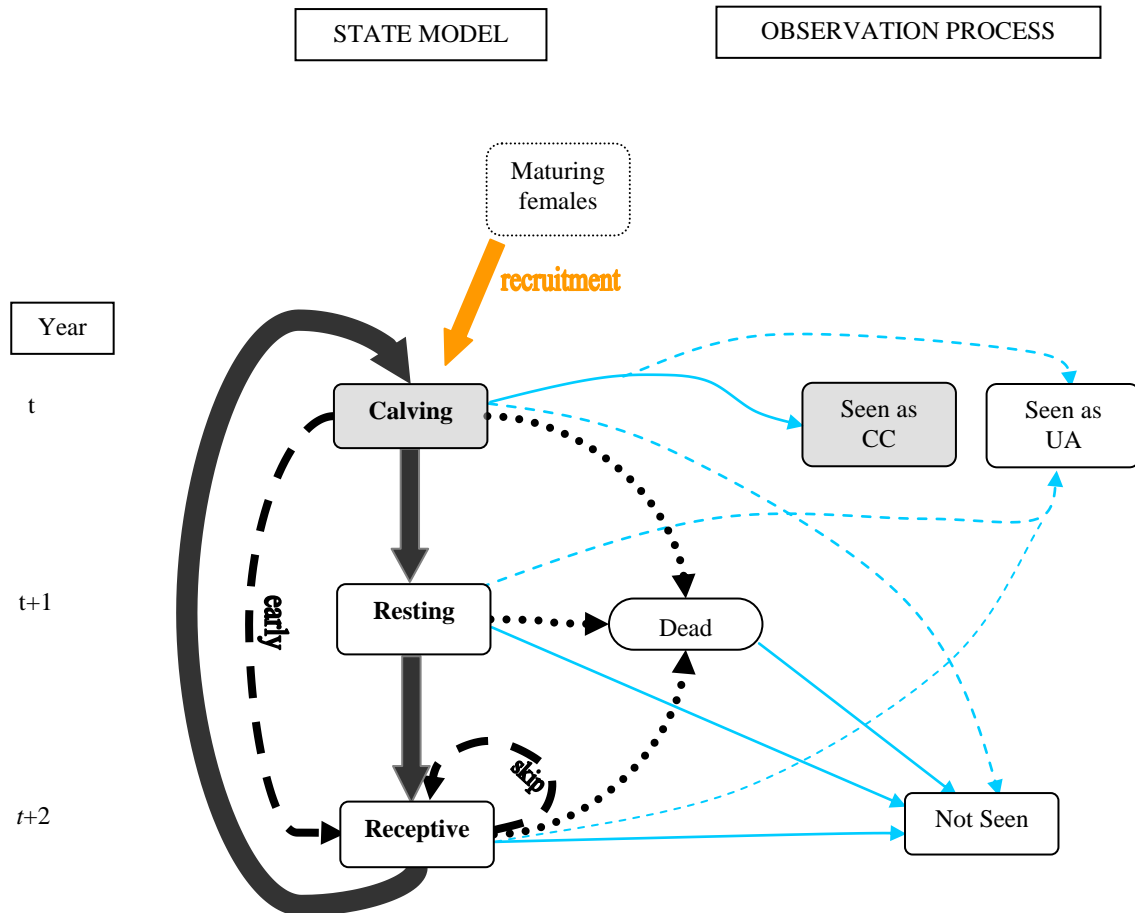
We assume that almost every calving female returns to a surveyed part of WASA. Most will be identified correctly as a cow-calf pair (CC), but a small proportion may be recorded as an unaccompanied adult (UA). Such a record need not be a 'mistake', for example if the calf is stillborn or dies before the survey.

True UAs (adult females without calves, i.e. in resting/receptive years) may also sometimes return to WASA, where they may or may not be seen. However, an UA will never be mis-recorded as a CC. If a female does not appear in the photo-ID record for a particular year, there are three possible reasons:

- (i) she may be calving but may not be photographed, at least not to sufficient quality;
- (ii) she may not be calving, and therefore either did not return, or did return but is not seen; or
- (iii) she may be dead.

If a female is recorded as an UA, then she is either in the resting/receptive phase and has simply decided to visit WASA that year, or she is calving but has been recorded as an UA. If she is recorded as CC, then she is definitely calving.

The same model is applicable to HoB. The chance of failing to record a CC in HoB is lower, because the survey period is much longer and there are more opportunities to obtain photographs. However, females that frequent HoB may still go unobserved in calving years if they calve somewhere else.



**Figure 2** Schematic model of the Australian southern right whales' breeding cycle ('State Model'). The 'Observation Process' shows how the whales' states are linked to what data may actually be observed and recorded. CC= cow/calf pair; UA=unaccompanied adult. Thick lines are indicative of the typical breeding cycle. Solid lines show what is normally happening, or is observed; dashed lines indicate lower probability events. Dotted lines represent death (or permanent emigration).

We note the importance of distinguishing between the *state* of a whale (Calving, Resting, Receptive, Dead, or Male), and the *observation* that may be made of it in any year (CC, UA, none). The observation made of a whale does not uniquely define its state, and the states follow a semi-regular biological rhythm that is not directly observed. These are the hallmarks of a multistate mark-recapture (MSMR) model (see, for example, Pradel, 2005; Conn and Cooch, 2009; Buchanan and Skalski, 2010).

Figure 2 is a schematic depiction of the model: the thick dark arrows represent the typical 3-year breeding cycle; dashed lines represent deviations from the norm, as described above; and death can occur from any given state (dotted lines). The right-hand part of Figure 2 shows the observation process (the actual data with which the model may be fitted) besides the state model.

Let  $t$  = year, and  $s$  = state (**B**, **R1**, **R2**, **D**), where **B** = Breeding (calving); **R1** = resting; **R2** = receptive; and **D** = died. The parameters of the model are:

#### Year-dependent

- $P[\text{skip} | t]$ : probability of remaining in **R2**, rather than going on to **B** (denoted by ‘*skip*’ in Figure 2). Together with the estimated variance of these probabilities, these are potentially of direct interest for monitoring because they relate to Antarctic conditions.
- $P[\text{obs} | s = \mathbf{B}, t]$ : probability that a calving female will appear in the photo-ID record.

#### Assumed time invariant

- $P[\text{early}]$ : probability of jumping straight from **B** to **R2**.
- $P[\text{dying}]$ : can die from any living state with equal probability.

#### And, if UA observations are included:

- $P[\text{obs} | s \neq \mathbf{B}]$ : although the absolute probabilities of being observed (included in the photo-ID record) if not in the breeding state are not individually identifiable, the relative probabilities ( $P[\text{obs} | s = \mathbf{R1}]$  and  $P[\text{obs} | s = \mathbf{R2}]$ ) are.  $P[\text{seen} | s = \mathbf{D}]$  is defined as zero.
- $P[\text{UA} | \text{CC}]$ : will a calving female be recorded as unaccompanied?



- $P[CC|UA]$ : will an unaccompanied adult be recorded as a calving female? This is considered very unlikely, and is set to zero for the scenarios considered in this report.

All observations conditional on state, and all transitions between states, are assumed to be independent Bernoulli (i.e. Binomial with  $n=1$ ) events, with probabilities as determined by the above model. Given those probabilities, the likelihood of the observations can be computed in closed form, using first-order Markovian assumptions (e.g. Schwarz *et al.*, 1993).

The year-dependent probabilities are estimated as random effects. This seemingly entails two extra parameters for the long-term mean and variance of skip probability (on a logit scale), and similarly for the probability of observation. Despite the extra parameters, the overall complexity, as defined by the effective degrees-of-freedom, is in fact greatly *reduced* with random-effect models, because the individual year-by-year probabilities are constrained not to show excessive variability.

The decision to use random-effects, rather than fully independent year-by-year ‘fixed effects’ (something of a misnomer for time-varying parameters) as has historically been more usual in mark-recapture, is partly borne of necessity: we discovered that the data are too sparse to allow reliable estimation of independent year-by-year effects. In addition, though, a random-effects framework provides a sound logical basis for simulating future datasets under different survey scenarios, whereby skip-breeding in the future is assumed to be about as common as it has been in the past (though this could be varied if desired).

In principle, the  $P[early]$  and  $P[dying]$  parameters could also be year-varying random effects. However, both phenomena are comparatively rare, and the datasets are still too short and with too few whales to allow estimation of the variability in events that are uncommon to begin with. See section 3.3 for more investigation of variability in mortality.

With respect to data selection, several minor variants can be considered. After experimentation, we chose to discard observations of a whale before its year of first calving (except for in estimating recruitment). Very little information is gained by including those observations, and by discarding them we avoided the substantial complexities of having to allow for males and immature females.

### 3.2: Estimation and software

To the best of our knowledge, there is no off-the-shelf software that fits random-effect MSMR models, and certainly none that also incorporates count data as per section 4, so we wrote bespoke software for this problem. We first estimate the random-effect variances by using Laplace approximation to integrate out all other parameters, and then form the posterior modes of those

parameters conditional on the estimated variances of the random-effects. Cooke *et al.* (2003) followed a similar route, although they used a different approximation that is somewhat less accurate in theory (though not necessarily in practice). Estimation is very quick, taking only about 60s even when count data are included, and numerical convergence diagnostics appear to be reliable. Goodness-of-fit diagnostics are discussed below.

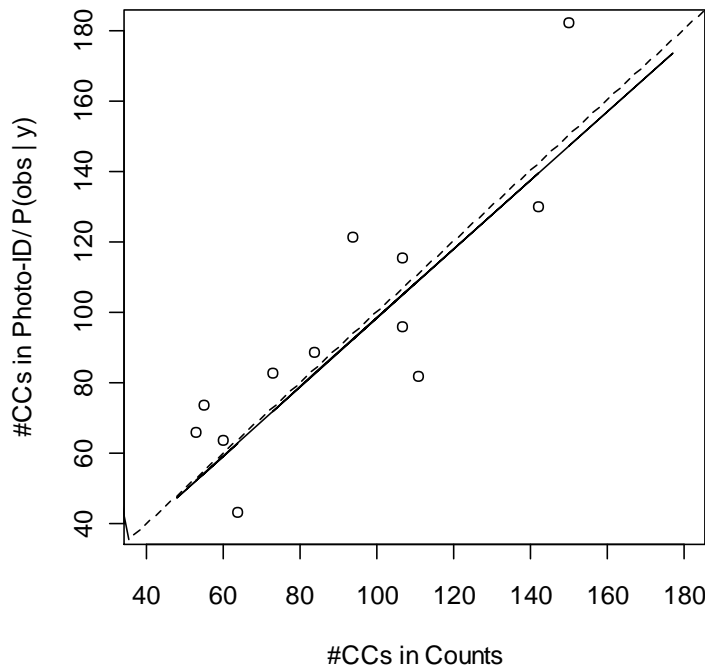
Approximate variances for the random and fixed effects (i.e. the probabilities) can be obtained simply from the Hessian of the 'likelihood' with respect to the non-variance parameters, including the random-effect prior terms. These are conditional on the point estimates of the random-effect parameters. For derived quantities with a strong nonlinear dependence on the parameters, such as population size, variance can be obtained by simulation, each time sampling the probabilities from a Gaussian distribution centered on the posterior mode and with the variance just described.

### 3.3 Diagnostics

In section 2, we assumed that the counts of cow-calf pairs were unbiased indices of the number of females breeding that year. But are they? If the counts really are close to a complete census of breeding females, then for each year,  $t$ , we should have:

$$\#CC \text{ in photos in year } t \sim \text{Bin}(\#CC \text{ count in year } t, P[\text{obs} | \mathbf{s} = \mathbf{B}, t] )$$

(neglecting for now any CCs classed as UAs). Consequently, a scatterplot of the number of CCs in the photo-IDs adjusted by the estimated observation probability for that year against the number of CCs from the counts should lie around the (0,1) straight line. This is confirmed in Figure 3. The fitted (solid) line lies slightly below equality (dashed line), but the difference is not statistically significant ( $p=0.42$ ).



**Figure 3** Number of cow-calf pairs from WASA photo-ID data against cow-calf pairs in WASA counts., for 1993-2006 The photo-ID counts in each year have been adjusted by the estimated observation probabilities from the multistate mark-recapture model. Dashed line is equality; solid line is fitted regression line.

A second diagnostic is whether the MSMR models correctly reproduce the *observed* intervals between calving (i.e. including the possibility that some calvings will go unobserved). It is hard to produce a formal goodness-of-fit statistic, because the effective degrees of freedom are not easy to define precisely. However, Table 1 shows that the model does seem reasonably consistent with the data, whether or not UA observations are used; the deviations between observed and expected counts in the 2-7 year-gap range are quite plausible for Poisson variability.

	Calving interval (years)												
	1	2	3	4	5	6	7	8	9	10	11	12	13
Obs	0	17	159	47	14	36	8	5	8	3	1	3	0
CC	0	16.6	165.9	37.2	14.9	32.9	13.7	5.9	6.2	3.0	1.1	1.1	0.4
CC+UA	0	18.1	164.8	35.6	14.4	33.6	13.1	5.7	6.2	2.9	1.0	1.1	0.4

**Table 1** Observed (Obs) calving intervals from WASA surveys, compared to expected intervals from models fitted to breeding females only (CC) and to breeding females and unaccompanied adults (CC+UA).

A third diagnostic is whether there is any evidence of systematic variation in mortality. With a MSMR model, it is possible to hindcast the probability distribution of the state of each animal in each year (conditional on the point estimates of population-wide probability parameters), using a ‘smoothing pass’ (the Viterbi algorithm). From this, an estimate of the ‘per animal mortality’ can be obtained by looking at the annual increment in probability of being dead, and this can be summed across animals to get an annual estimate. The results for the model without UA observations are given in Table 2 below:

	Year											
	1994	1995	1996	1997	1998	2000	2001	2002	2003	2004	2005	2006
%death rate	3.2	3.6	4.0	3.9	3.2	2.9	2.9	2.9	3.0	3.0	3.0	3.0

**Table 2** Estimated death rates (given as %) for each year for model fitted to calving females.

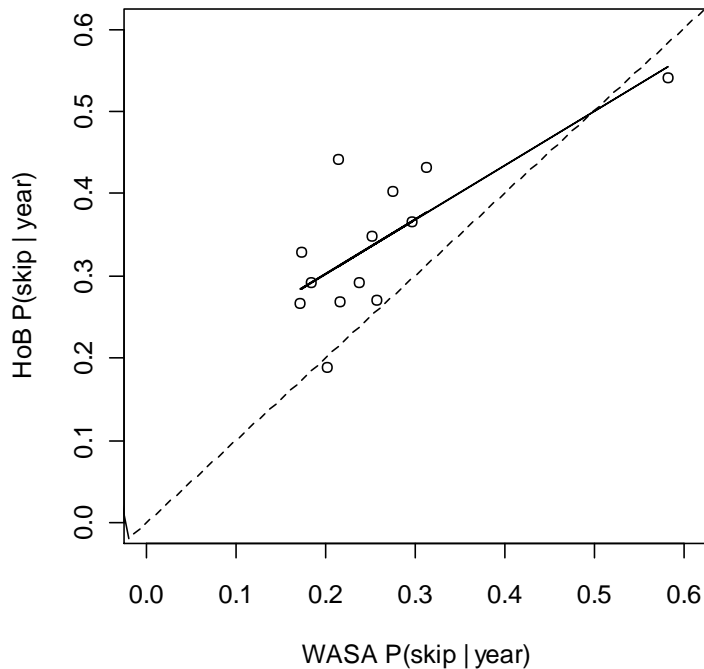
The variation is very low, just 2.9-4.0%. Like most diagnostics based on some notion of residuals, this analysis is not ‘looking for trouble’ and its sensitivity has not been tested. Nevertheless, the mortality diagnostic could be useful in future analyses for detecting substantial mortality increases, at least over a period of years; it will always be rather difficult to pin down individual mortality events with animals that are only seen every third year.

### 3.4: Comparison between sites

As already noted, there do seem to be substantial differences between the WASA and HoB population dynamics (Burnell, 2008; Bannister 2009). This is apparent just from inspecting the frequency distribution of gaps between observed calves. The HoB data shows relatively fewer 2-year, and relatively more 4-year, gaps than the WASA data (see the rows labelled ‘Obs’ in Table 3). This is despite the probability-of-observation being higher in HoB, which would actually tend to shorten the average observed gap in HoB compared to WASA. The MSMR model results tell the same story: a lower P(early) and a higher mean P(skip) in HoB. The longer gaps imply a slower rate-of-increase for HoB compared to WASA, and indeed this is what Burnell (2008) found (for cow-calf pairs a rate of increase of 3.6-5.4% p.a. from the HoB count data) cf. 7.5% p.a. for cow-calf pairs from the aerial counts (Bannister, 2009).

		Calving interval (years)												
		1	2	3	4	5	6	7	8	9	10	11	12	13
WASA	Obs	0	17	159	47	14	36	8	5	8	3	1	3	0
	Exp	0	16.6	165.9	37.2	14.9	32.9	13.7	5.9	6.2	3.0	1.1	1.1	0.4
HoB	Obs	0	5	147	52	17	14	9	1	4	0	0	0	0
	Exp	0	5.4	176.6	55.3	16.9	18.5	10.4	4.3	3.0	1.8	0.8	0.5	0.2

**Table 3:** Observed and expected calving intervals from models fitted (separately) to CC photo-ID data from the WASA aerial surveys and HoB land-based survey.



**Figure 4:** Estimated skip probabilities from WASA and HoB photo-ID data, estimated using breeding females (CCs). There was an anomalously low count in 2003, with consequential high skip probability as indicated. Solid line is fitted linear regression line; dashed line is equality. Note that correlation between the two sets of estimates remains if the 2003 ‘outlier’ is excluded from the regression.

Given the evident differences between the population dynamics, it does not make sense to pursue the original project objective of a joint model that infers a *single*  $P(\text{skip})$  each year that can apply both to WASA and to the HoB (with different observation probabilities between sites and years, of course). It is still interesting to see how far HoB and WASA do tell the same story about skip-breeding though, and Figure 4 shows a scatterplot of point estimates of skip probabilities from models fitted separately to each dataset. Although the mean skip probabilities are different, there is a clear positive relationship, even after removing the 2003 ‘outlier’. (A formal test is not appropriate here, or at least not a simple test, since each set of estimates is internally highly correlated.) This suggests that the HoB photo-ID data may still have some predictive power for WASA, but that the nature of the ‘calibration’ would need to be estimated. A more complicated joint model could be developed along those lines. However, the two data series may not yet be long enough to establish a reliable calibration, and of course the past may not be a guide to the far future since the HoB sub-population seems to be following a different trajectory to WASA in general. It seems prudent to first get a better

understanding of why the dynamics are different, before developing a survey strategy that simply assumes the difference will remain constant.

#### 4: EXTENDING THE MSMR MODEL TO INCORPORATE COUNT DATA

Taking the CC count as a census, variability arises partly from year-skipping and breeding cycles, but also from fluctuations in the number of newly-recruited females. Since we only start the observation history of each whale from its first observed calving, we can for convenience also define ‘recruitment’ as pertaining to the first year that a whale calves (a slightly different definition from that of Cooke *et al.*, 2003, but equally driven by expediency). Given the parameters of skip-breeding, early-breeding, and mortality, and since the counts are assumed to be without significant error, the actual recruitment in each year is determined exactly, provided the count data series has no gaps; the recruitments are not really parameters to be estimated, in contrast for example, to the skip-breeding probabilities. However, the variability in recruitment can be inferred post hoc, simply from the time series of inferred recruitments. Recruitment variability is a key parameter for predicting the precision of a future survey strategy that does have gaps, and therefore unknown recruitments, and therefore uncertain population sizes.

To estimate recruitment variability, some kind of model for expected recruitment is required. Whereas it is possible to construct a fairly well-founded model for the breeding cycle, it is not really possible to do this for recruitment. With a maturity period of ~10 years, and an initial birth cohort size that must have depended on unknown abundance several years before the surveys began, an explicit model is impossible, and something more empirical is required. In choosing the model for recruitment, it is also important not to inadvertently constrain future recruitments to follow the trend seen in the past. The whole point of a monitoring strategy is that the future might not be exactly like the past - otherwise we could simply infer population size using a 7.5% rate of increase.

In this respect, a fully Bayesian analysis (which might be suggested by the random-effects framework we have used) could be problematic; the data so far give no reason to suggest anything other than an exponential rate-of-increase with low variability, but it would be disastrous to hard-wire that trend and variability into the estimation part of a monitoring strategy. It would be perverse not to have a ‘weak default’ for a trend in recruitment, but it is important that the recruitment variability in future should be assumed fairly high, so that the monitoring has a chance to be responsive to the data.

The model we used for expected recruitment used the proportion of ‘new’ photo-IDs of CCs in any given year. Most of the newly-identified animals are new recruits, but a proportion of them would have been whales that recruited

earlier but were not observed. Then, in order to incorporate the count data into the photo-ID model, it is simply necessary to include extra parameters for recruitment in unsurveyed years, and to compute the recruitment in all surveyed years so that the number of counted CCs exactly matches the true abundance. An additional term in the likelihood is computed to allow for variability in recruitment, based on the following:

$$\# \text{New CC in photos in year } t \sim \text{Bin}(\# \text{CC count in year } t, p_t^{\text{new}}),$$

where  $p_t^{\text{new}}$  is defined as the conditional probability of a whale being photo-identified before  $t$  as a CC, given that it was identified in year  $t$  as a CC.

In principle, it is possible to estimate ‘recruitment variability’ as another random-effect variance, in the same way as for the probabilities of skip-breeding and observation. However, the dataset is just too small at present for this to work with any degree of reliability.

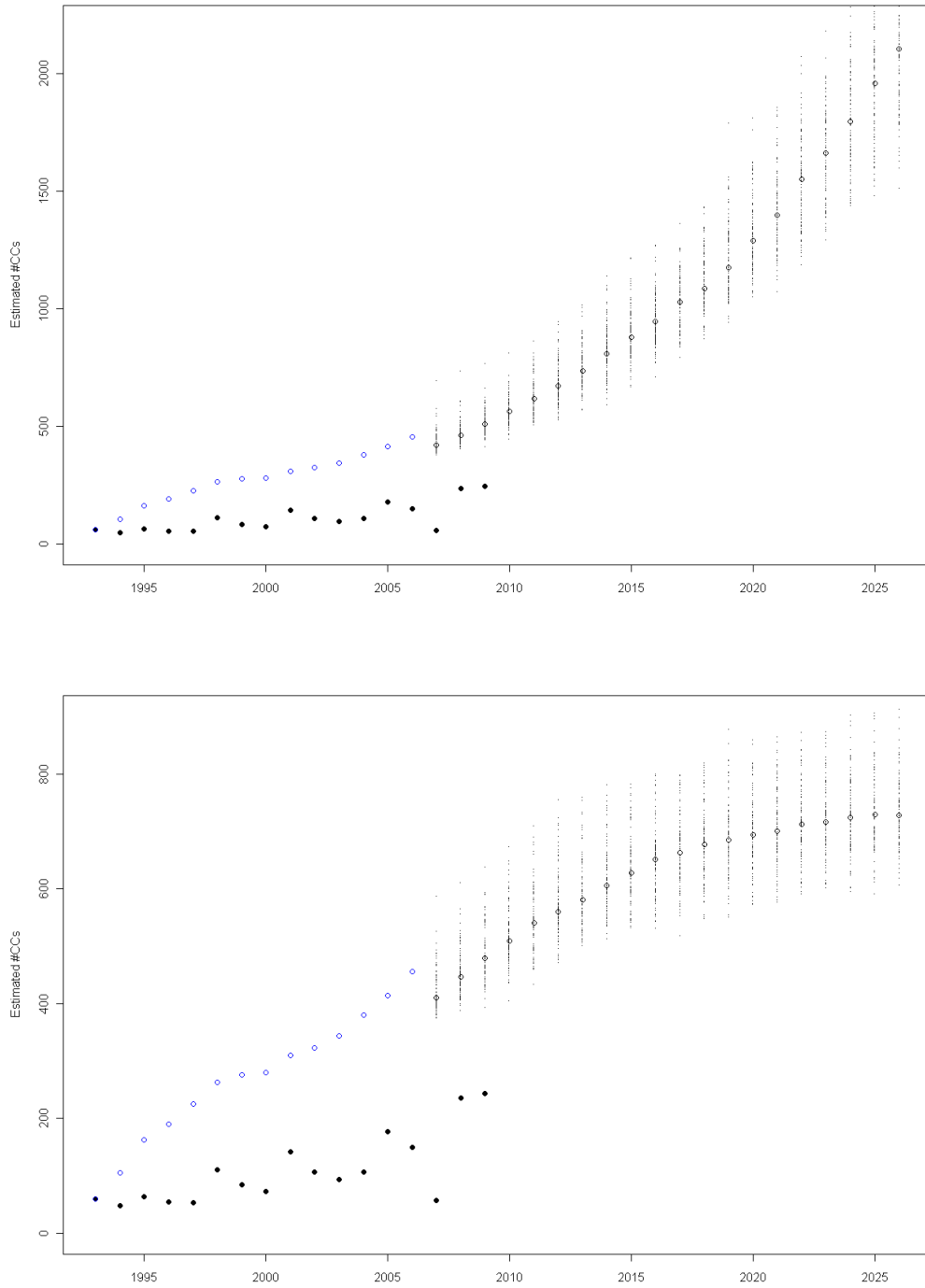
For the purposes of this project, our recruitment model seemed to be giving reasonable outcomes - it allowed sensible extrapolation of the data into the future with realistic variability. Thus the simulation results should not provide overly-optimistic estimator precision.

## 5: SIMULATIONS FOR FUTURE MONITORING

In order to consider future monitoring strategies, we generated extensions of the existing WASA count and photo-ID data, using parameter estimates from the joint photo-ID and count WASA model. For the fixed effects ( $P[\text{early}]$ ,  $P[\text{dying}]$  and  $P[\text{obs} | \mathbf{s} \neq \mathbf{B}]$ ), the point estimates from the model are used. Those things that are random effects - the skip probabilities, the observation probabilities and recruitment - are drawn annually from distributions with appropriate means (point estimates) and variances.

Mass-mortality events were not explicitly accounted for in the simulations for two reasons: a sudden and extensive mortality event would be immediately evident from the data; and a gradual increase in mortality over a prolonged period should be detected by the mortality diagnostic (see section 3.3). Considering reduced recruitment to be the more likely scenario, we examined two possible future population trajectories in the simulations: a continuation of the rate of increase similar to that seen in the WASA surveys to date, and reduced recruitment leading to annual rates of increase of only about 5%.

Typical simulated population trajectories from the two levels of recruitment are shown in Figure 5.



**Figure 5:** Sample population trajectories for different levels of rate of increase beyond 2006. Top panel = ROI about 9% p.a.; bottom panel = average ROI about 5% p.a. but flattening off to 3% p.a. by the end of the series. Circles are estimated numbers of CCs - blue circles are fitted to data; black circles to simulated data, with dots showing the true variation in numbers of CCs in the simulations. Filled black circles are CC counts from the aerial surveys. Note the low count in 2007.



For each of the two trajectories, we simulated 100 replicates of the following potential monitoring strategies:

- 1) Survey annually
- 2) Survey every other year
- 3) Skip one year, survey next two; repeat
- 4) Skip two years, survey next three; repeat
- 5) Skip two years, survey next four; repeat
- 6) Skip two years, survey next five; repeat

The point of looking at strategies 4-6 was to see what could be gained from surveying in consecutive years which encompassed different proportions of the breeding cycle. For example, does four years of consecutive surveying allow us to 'pin down' sufficiently well the skip-probabilities (from the photo-ID) that the counts (with their cohort structure) can be interpreted unambiguously?

Strategy 1 is 'continue as is', i.e. surveying annually and was examined to show the baseline CV in abundance that could be obtained from this regime, given full use of the photo-ID and count data.

Strategy 2 was investigated as a pragmatic reduction-in-effort approach, which in the absence of cohort structure may have appeared reasonable. Strategy 3 was selected as a regime which *a priori* might have been expected to perform rather poorly, given the 3-year breeding cycle.

Projections were for 20-21 years from 2006, depending on the monitoring strategy. For examining the effect of the different survey regimes on monitoring population size, the CVs (coefficients of variation) of estimated abundance (of breeding females) were estimated using the simple measure:

$$\frac{\sqrt{\text{var}(\text{true \# CCs} - \text{estimated \# CCs})}}{\text{mean}(\text{estimated \# CCs})}$$

(since the true number of CC pairs is known from the simulations). Results are shown in Table 4. Note that these results use *all* the data to estimate each year, so they are the most up-to-date CVs.

		Year (2000+)																				
		07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
rct level as is	1	2	2	1	2	2	2	2	2	2	2	2	2	1	1	1	1	2	1	1	2	
	2	9	9	14	10	13	10	12	9	12	10	14	10	10	9	14	10	15	11	14	11	
	3	9	8	7	12	10	8	12	11	9	13	11	9	13	12	10	14	12	10	15	12	11
	4	8	12	11	8	3	11	18	16	10	4	10	16	14	9	4	12	17	15	9	3	
	5	8	13	12	8	4	2	12	16	14	9	4	2	13	18	16	10	4	2	12	18	
	6	7	13	12	8	4	2	2	12	18	16	11	4	2	2	10	16	14	10	4	2	2
reduced rct	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	
	2	8	7	11	9	12	9	11	9	10	8	10	8	9	8	9	8	8	7	10	10	
	3	8	8	7	11	9	8	11	10	9	10	10	8	10	10	9	11	11	10	10	10	
	4	7	10	10	7	4	8	12	12	8	4	7	9	9	6	4	7	9	9	7	5	
	5	8	13	12	9	4	3	8	12	11	8	3	2	6	9	9	6	3	3	6	8	
	6	8	11	10	8	4	3	3	9	12	12	8	4	3	3	6	9	8	6	3	2	2

**Table 4:** Estimated percentage coefficients of variance in abundance of CC pairs under different monitoring strategies 1-6, and for two levels of recruitment: continuing approximately as is, and reduced recruitment (an average of 5% p.a., flattening to about 3% p.a. by the end of the series). Greyed out numbers represent no survey under the given regime in that year.

### 5.1 Interpreting these results

What these results immediately tell us is the following:

- Under either existing recruitment levels, or under reduced recruitment (actually simulated to begin occurring in 2007 since for the present analyses the photo-ID data from WASA had been processed up to 2006), high precision, with CVs of the order of 1-2% could be expected with annual surveys.
- With either ‘pragmatic’ reduction in effort, surveying every other year (strategy 2), or with the deliberately-chosen poor strategy of surveying two out of every three years for a population with a three-year (but fluctuating) breeding cycle (strategy 3), we are never going to achieve a precise estimate. For a given year, CVs range from 7-15% and 7-12% depending on the recruitment level. Average CVs from strategy 2 were fractionally more precise than those from strategy 3 for the existing recruitment case, but the reverse was true for the reduced recruitment level. Certainly there is no evidence that an ill-advised survey monitoring level of 67% is any better than a level of 50%.
- Surveying with gaps of two years with consecutive surveys for 3, 4 and 5 years (equivalent to 60%, 67% and 71% effort; strategies 4-6) are each more reasonable strategies than 2 and 3.
- Precision in the estimates from strategies 4-6 is cyclical. Loss of precision is evident during the gap-years (particularly the second of the two), and then there is a gradual increase in precision during the blocks of survey years, of the order of 2-3% for the fourth, or fourth and fifth, year(s) of strategies 5 and 6 respectively.

Over the long term, then, badly-chosen gappy strategies are bad, but a well-chosen gappy strategy is still eventually able to deliver a *very* precise abundance estimate. However, the price paid is the time taken to reach that precision. With a 2-years-off, 4-years-on strategy, it takes six years to recover full precision. Whether that is acceptable or not, depends on the question of interest. In terms of truly long-term population dynamics, and for a population like this where the doubling time is about nine years, perhaps a six-year delay is acceptable. However, if there is interest in monitoring for more sudden changes, then a six-year wait may well be too long, particularly if there is a need to consider a management response. We do not attempt here to second-guess the ‘acceptability’ criterion.

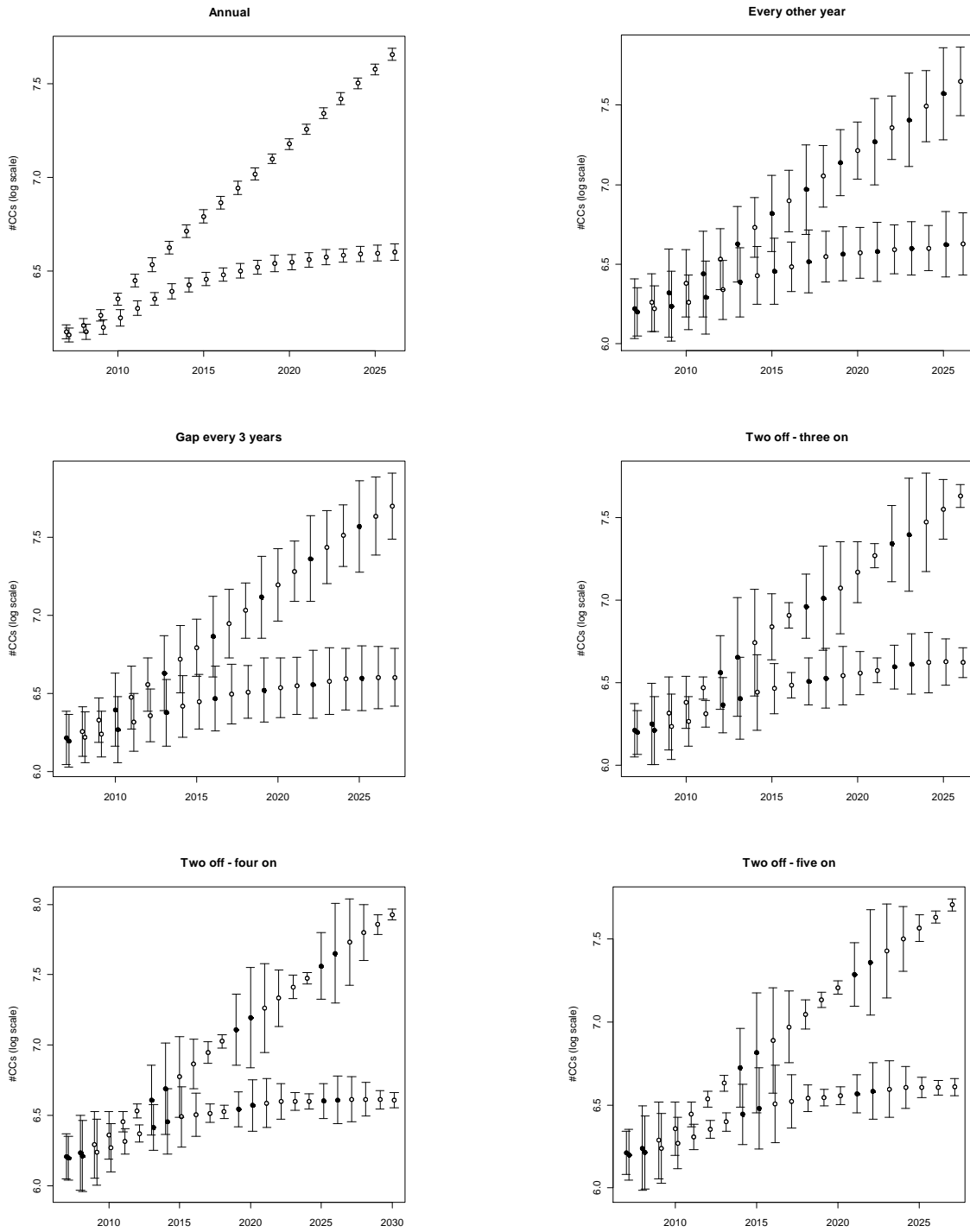
How do these results help in designing future monitoring schemes? There are various questions one might ask, most of which are not covered directly in this report (but which the software may still be of use in answering). We focus on loss of precision (and subsequent failure to detect changes in abundance) by *not* surveying annually.

Consider the two (true) population trajectories of Figure 5. Twenty years into the ‘future’ (i.e. beyond 2006), the population would be expected to total around 2100 adult females in the ‘base case’ scenario (representing a Rate of Increase similar to that estimated from the existing data), or about 730 adult females given the decline in recruitment as shown by the trajectory in the lower panel.

For both trajectories and under the six various survey strategies described above, we computed the estimated mean population sizes for each year and their associated CVs. Log-based confidence intervals were calculated; these are shown in Figure 6.

The first plot shows that with annual surveys we are able to detect a significant decline in recruitment after four years (for the levels of recruitment assumed). This compares with 10 years for biennial surveys, and 11 years for the 1-year-off, 2-years-on strategy.

For the strategies with increasing number of consecutive survey years, the retrospective advantage of incorporating the photo-ID data becomes clearer. For strategy 4 (2-years-off, 3-years-on), there is a chance that a significant decline might be detected after just 5 years, i.e. at the end of the 3 years of consecutive survey. However, by not surveying for the subsequent two years, then we would not be certain that this is a real decline rather than some quirky low-recruitment year as the confidence intervals in the gap-years are so much wider (and hence are not able to detect significant changes despite the increasingly divergent trajectories). Strategies 5 and 6 do rather better in this regard; both detecting a significant decline after 5 years, with confirmation of this in the fourth (and for strategy 6, the fifth) consecutive year(s) of survey.



**Figure 6:** Estimated log(adult female abundance) and confidence intervals from fitting the joint photo-ID and count model to the simulated data, under six survey scenarios and two possible population trajectories. Open circles are mean point estimates from years with surveys; filled circles are from non-survey years. The reduced-recruitment estimates have been slightly offset for ease of readability when the confidence intervals overlap.

## 5.2 Software for future monitoring of Southern right whales

As part of this project, it became necessary to write extensive bespoke software in order to fit the joint model to the photo-ID and count data from WASA. We are happy to make the software freely available for use by those interested, either to consider additional possible survey strategies not reported here, or for use in the future. The package, named 'srw', is written (and reasonably well-documented!) in R (R Development Core Team, 2011), with underlying Pascal, and is available from Mark Bravington (mark.bravington@csiro.au). The estimation algorithm relies on Automatic Differentiation (AD; Griewank, 2008) to handle the random effects, as described in Skaug and Fournier (2006), although our software implementation is different and makes use of the TAPENADE AD engine (INRIA Tropics team, unpublished).

## 6: CONCLUSIONS

At the outset of this project, it was envisaged that future monitoring strategies would focus on how the Head of Bight data (presumed not to have gaps) might be used to complement the counts from a 'gappy' aerial survey in WASA. Provided that the population dynamics at the HoB were sufficiently representative of those in WASA, the photo-ID information from HoB might provide enough information about skip-years to largely eliminate the precision lost in gap-years from WASA. In the event, though, simple data summaries of the observed calving intervals from the Head of Bight site compared with the calving intervals for the whole WASA area indicated substantial difference in population dynamics between the two surveys. Specifically, the HoB had fewer 2-year calving intervals (as inferred from, and confirmed by the estimate of  $P[\text{early}]$ ), and more 4-year intervals ( $P[\text{skip}|t]$ ), than WASA. It was not clear why there is such a difference; perhaps the HoB animals are older and less fecund, but doubtless there are other plausible explanations. In any case, until there is a good explanation and a clear rationale, it does not seem advisable to patch up gaps in WASA coverage using HoB photo-ID data. This gives less scope for ingenious efficiency in survey design, but our WASA-only model still shows how much precision can be gained by using the photo-ID data as well as the counts.

We have experimented with a number of different ways to estimate variance from the model, rather than resorting to using the simulated 'truth', but to date these have proved unsuccessful. Therefore we have not examined in detail estimation of skip-probability variances - as pertaining to detection of environmental linkages. It does however seem clear that if detection of environmental linkages is considered an important aim of future surveys, then annual surveys would be much better than any reduced-frequency strategy.

In this report, the future monitoring strategies have focused on using the existing data to predict what levels of estimator precision might be obtained in the future. No consideration has been given at this point to survey logistics. Two specific logistic details that need to be contemplated are: (i) would a more 'gappy' survey schedule affect continuity of logistic expertise and resources; and (ii) in the event that the population continues to increase (at the current WASA rate, doubling every nine years or so), will the same proportion of cow-calf pairs still be able to be successfully photographed, or is there an absolute limit? If the proportion of whales photo-identified were to decline, then the observation probabilities would decrease, with consequential loss in estimator precision.

Our results in section 5 suggest that surveys should continue annually, at least for the next few years. The models we have developed incorporating count and photo-ID data give more stable estimates of the number of adult females in the population each year than those which use the count data alone, and should be able to detect poor-breeding years. They serve to highlight that the aerial survey series, for which photo-ID data from 1993-2006, and count data from 1993-2009 (now 2010) were available, is still relatively short for a population exhibiting such a strong cohort structure. The population dynamics model (Cooke *et al.*, 2001, 2003) and resulting inferences relating these dynamics to environmental signals in the Southern Ocean were based on a substantially longer, 30-year data series on South American southern right whales, with formation of hypotheses supplemented by data series of similar length on other Southern Ocean predators, such as gentoo penguins. The WASA series is currently too short to estimate recruitment variability in a sophisticated manner; we recommend that the photo-ID data from 2007-2010 are incorporated into the database with high priority - these would help considerably for this aspect of the estimation.

## REFERENCES

- Bannister, J.L. 1998. Aerial survey of southern right whales off Western and South Australia, 1997 calving season. Report on work undertaken to 31 January 1998 to Biodiversity Group, Environment Australia (unpublished) 18pp.
- Bannister, J.L. 2002. Southern right whale aerial survey and photoidentification, southern Australia coast, 2001 calving season. Final Report to Environment Australia, (unpublished) 18pp.
- Bannister, J.L. 2009. Southern right whale aerial survey, southern Australian coast, August 2009. Final Report on work funded by The Island Foundation, Massachusetts, U.S.A. (unpublished) 16pp.
- Bannister J L, 2011. Southern Right Whale Aerial Survey, Southern Australian Coast, August 2010. Final Report on work funded by the Australian Government through the Australian Marine Mammal Centre (unpublished) 16pp.

- Buchanan, R.A. and Skalski, J.R. 2010. Using multistate mark-recapture methods to model adult salmonid migration in an industrialized river. *Ecological Modelling* 221: 582-589.
- Burnell, S.R. 2008. Estimates of demographic parameters of southern right whales off Australia. Paper SC/60/BRG12 presented to the Scientific Committee of the International Whaling Commission, Santiago, June 2008 (unpublished) 8pp.
- Conn, P.B. and Cooch, E.G., 2009. Multistate capture-recapture analysis under imperfect state observation: an application to disease models. *J. Applied Ecology* 46: 486-492.
- Cooke, J.G., V. Rowntree, and R. Payne. 2001. Estimates of demographic parameters for southern right whales (*Eubalaena australis*) observed off Península Valdés, Argentina. *J. Cetacean Research and Management*, Special issue 2:125-132.
- Cooke, J.G., V. Rowntree, and R. Payne. 2003. Analysis of inter-annual variation in reproductive success of South Atlantic right whales (*Eubalaena australis*) from photo-identifications of calving females observed off Península Valdés, Argentina, during 1971-2000. Paper SC/55/023 presented to IWC Scientific Committee, Berlin, June 2003 (unpublished) 16pp.
- Griewank, A.W. 2000. *Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation*. *Frontiers in Applied Mathematics No. 19*, Society for Industrial and Applied Mathematics, Philadelphia.
- INRIA TROPICS Team. TAPENADE v2.1 software for Automatic Differentiation. Available from <http://www-sop.inria.fr/tropics/> .
- Leaper, R., Cooke, J., Trathan, P., Reid, K., Rowntree, V. and Payne, R. 2006. Global climate drives southern right whale (*Eubalaena australis*) population dynamics. *Biology Letters*, 2: 289-92.
- Pirzl, R., Thiele, D., Bannister, J. and Burnell, S. 2008. ENSO and SAM affect reproductive output in southern right whales. Attachment A of Report to the Department of Environment, Water, Heritage and the Arts (unpublished) 25pp.
- Pradel, R. 2005. Multievent: an extension of multistate capture-recapture models to uncertain states. *Biometrics*, 61: 442-447.
- R Development Core Team. 2011. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-90051-07-0, URL <http://www.R-project.org>.
- Schwarz, C. G., Schweigert, J. F., and Arnason, A. N. 1993. Estimating migration rates using tag recovery data. *Biometrics* 59: 291-318.
- Skaug, H.J. and Fournier D. 2006. Automatic Approximation of the Marginal Likelihood in non-Gaussian Hierarchical Models. *Computational Statistics and Data Analysis* 51: 699-709.